

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:25:00 ; Search time 97 Seconds
(without alignments)
3251.750 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 5034
Sequence: 1 MSSGLWSQEKVTSYWEERI.....RLLCDAYMCYQSPMSLYK 949

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4983	99.0	960	16	US-10-788-792-250
2	4971.5	98.8	956	16	US-09-851-673-4
3	4971.5	98.8	956	16	US-10-755-889-490
4	3862	76.7	731	16	US-10-921-707-9
5	1275	25.3	261	15	US-10-264-237-1609
6	1173	23.3	238	15	US-10-264-237-1610
7	597	11.9	113	9	US-09-864-761-34675
8	146	2.9	354	13	US-10-087-192-120
9	143	2.8	547	15	US-10-108-260A-2751
10	143	2.8	547	15	US-10-275-595A-13
11	135.5	2.7	306	15	US-10-104-047-2990
12	135.5	2.7	721	11	US-09-764-875-857
13	135.5	2.7	2273	15	US-10-282-122A-66115

14	134.5	2.7	721	9	US-09-764-868-731	Sequence 731, Appl
15	134.5	2.7	721	11	US-09-764-875-1140	Sequence 1140, Ap
16	134.5	2.7	816	16	US-10-437-963-132799	Sequence 132799,
17	134	2.7	307	14	US-10-106-698-5606	Sequence 5606, Ap
18	133.5	2.7	1319	16	US-10-408-765A-343	Sequence 343, App
19	132.5	2.6	439	13	US-10-087-192-117	Sequence 117, App
20	130	2.6	717	15	US-10-369-493-22287	Sequence 22287, A
21	129.5	2.6	1281	16	US-10-363-829-373	Sequence 373, App
22	129.5	2.6	3298	14	US-10-160-758-16	Sequence 16, Appl
23	129.5	2.6	3298	14	US-10-174-677-8	Sequence 8, Appli
24	129.5	2.6	3298	15	US-10-120-801-51	Sequence 51, Appl
25	129.5	2.6	3298	15	US-10-210-172-50	Sequence 50, Appl
26	127.5	2.5	3217	15	US-10-311-623-8	Sequence 8, Appli
27	127	2.5	803	16	US-10-437-963-132702	Sequence 132702,
28	126.5	2.5	808	16	US-10-437-963-132681	Sequence 132681,
29	124	2.5	1474	14	US-10-225-567A-522	Sequence 522, App
30	124	2.5	1474	15	US-10-292-798-914	Sequence 914, App
31	124	2.5	2621	16	US-10-437-963-122168	Sequence 122168,
32	123.5	2.5	814	16	US-10-437-963-185098	Sequence 185098,
33	122.5	2.4	803	14	US-10-349-436-33	Sequence 33, Appl
34	122.5	2.4	892	15	US-10-276-774-1800	Sequence 1800, Ap
35	122.5	2.4	907	17	US-10-491-213-8	Sequence 8, Appli
36	122.5	2.4	1120	16	US-10-437-963-152821	Sequence 152821,
37	122.5	2.4	1282	16	US-10-437-963-110654	Sequence 110654,
38	122.5	2.4	1953	15	US-10-369-493-1945	Sequence 1945, Ap
39	122	2.4	2263	16	US-10-408-765A-2231	Sequence 2231, Ap
40	121	2.4	995	16	US-10-437-963-105117	Sequence 105117,
41	120.5	2.4	867	15	US-10-282-122A-57767	Sequence 57767, A
42	120.5	2.4	1033	9	US-09-888-615-75	Sequence 75, Appl
43	120.5	2.4	1109	15	US-10-425-114-72939	Sequence 72939, A
44	120.5	2.4	1471	10	US-09-998-027-4	Sequence 4, Appli
45	120.5	2.4	1471	14	US-10-165-099-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-788-792-250
; Sequence 250, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 250
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-788-792-250

Query Match	99.0%	Score 4983	DB 16	Length 960
Best Local Similarity	99.5%	Pred. No. 0		
Mismatches	948	Conservative	0	Mismatches 1; Indels 4; Gaps 4;
Qy	1	MSSGLWSQEKVTSYWEERIFYLLOQCSVTDKOTKLLKVPKSGISGOYIDRSVGHRI	60	
Db	8	MSSGLWSQEKVTSYWEERIFYLLOQCSVTDKOTKLLKVPKSGISGOYIDRSVGHRI	67	
Qy	61	PSAKGKNQIGLKILEOPHAFVDE-DVVINEKFTELLALITNCEERFSLFKNRRLS	119	
Db	68	PSAKGKNQIGLKILEOPHAFVDEKDVINEKFTELLALITNCEERFSLFKNRRLS	127	
Qy	120	KGLQIDVCGPVKQLRSGEKFPVGRFPLAERTVSGIPFGVELLEEGRGQFTDGV	179	

Db 128 KGLQIDVGPVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRCQGFDTGV 187
Qy 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRVSILK 238
Db 188 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRVSILK 247
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHN 297
Db 248 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGVQVLCFSFACVESTILLHN 307
Qy 298 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNPKATGTSDDGNNR-RSELFTYTLNGSSV 356
Db 308 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNPKATGTSDDGNNRSELFTYTLNGSSV 367
Qy 357 DSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRPHSLP 416
Db 368 DSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRPHSLP 427
Qy 417 TKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSLAEVKNPP 476
Db 428 TKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSLAEVKNPP 487
Qy 477 FYGVIRWIGOPPGNLEVLAGELEDEACAGTGTFRGTRYFTCALKKALFVKLSKCRPDS 536
Db 488 FYGVIRWIGOPPGNLEVLAGELEDEACAGTGTFRGTRYFTCALKKALFVKLSKCRPDS 547
Qy 537 RFASLQPVNSQIRCNLSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
Db 548 RFASLQPVNSQIRCNLSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 607
Qy 597 DSTLFCLFAPSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIMKL 656
Db 608 DSTLFCLFAPSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIMKL 667
Qy 657 KILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
Db 668 KILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 727
Qy 717 KVGVPITIQOLLEWSFINSNLKFAEAPSCLLIQMPRFKGFPLFKKIFPSELEINITLLED 776
Db 728 KVGVPITIQOLLEWSFINSNLKFAEAPSCLLIQMPRFKGFPLFKKIFPSELEINITLLED 787
Qy 777 TPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTKNTQVHLHPKRLNKNPVSILPKD 836
Db 788 TPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTKNTQVHLHPKRLNKNPVSILPKD 847
Qy 837 LPDWDNRHGICPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNGFNI 896
Db 848 LPDWDNRHGICPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNGFNI 907
Qy 897 PQVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMVQSPMTSLYK 949
Db 908 PQVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMVQSPMTSLYK 960

RESULT 2

US-09-851-673-4
; Sequence 4, Application US/09851673
; Publication No. US20030165985A1
; GENERAL INFORMATION:
; APPLICANT: Derry, Jonathan
; APPLICANT: Fenslow, William
; APPLICANT: Dougall, William
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
; FILE REFERENCE: 3198
; CURRENT APPLICATION NUMBER: US/09/851, 673
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 956
; TYPE: PR1
; ORGANISM: Homo sapiens

late

US-09-851-673-4

Query Match 98.8%; Score 4971.5; DB 10; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;
Qy 1 MSSGLMSQKVTSPYWEERIFYLLQECSTVDKQTKLLKVPKGSIGQYIQDRSVGHSHRI 60
Db 1 MSSGLMSQKVTSPYWEERIFYLLQECSTVDKQTKLLKVPKGSIGQYIQDRSVGHSHRI 60
Qy 61 PSAGKKNQIUGLKILQPHAVLFVDE-DVVEINEKFTTELLAINTCERESLFPKNRRLS 119
Db 61 PSAGKKNQIUGLKILQPHAVLFVDEKDVVEINEKFTTELLAINTCERESLFPKNRRLS 120
Qy 120 KGLQIDVGPVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVELLEBGRGQGFDTGV 179
Db 121 KGLQIDVGPVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVELLEBGRGQGFDTGV 180
Qy 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRVSILK 238
Db 181 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDTMQVLPPLPINSRVSILK 240
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHN 297
Db 241 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDGRPDGVQVLCFSFACVESTILLHN 300
Qy 298 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNPKATGTSDDGNNR-RSELFTYTLNG 353
Db 301 DIIPESVTOERRPPKLAFAFMSRGVGDGSSSHNPKATGTSDDGNNRSELFTYTLNG 360
Qy 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRPHSLP 413
Db 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRPHSLP 420
Qy 414 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSLAEVKE 473
Db 421 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGLEVGSLAEVKE 480
Qy 474 NPFFYGVIRWIGOPPGNLEVLAGELEDEACAGTGTFRGTRYFTCALKKALFVKLSKCR 533
Db 481 NPFFYGVIRWIGOPPGNLEVLAGELEDEACAGTGTFRGTRYFTCALKKALFVKLSKCR 540
Qy 534 PDSRFASLQPVNSQIRCNLSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 593
Db 541 PDSRFASLQPVNSQIRCNLSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 600
Qy 594 CYLDSTLFCFLFAPSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIM 653
Db 601 CYLDSTLFCFLFAPSSVLDTVLLRPKEKNVVEYSETQELLRTTEIVNPLRIYGVVCATKIM 660
Qy 654 KLRKILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVQDCYFYQIFME 713
Db 661 KLRKILEKVEAASGFTSEKDPPEEFNLIFPHILRVEPLLKIRSAQKQVQDCYFYQIFME 720
Qy 714 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSCLLIQMPRFKGFPLFKKIFPSELEINITDL 773
Db 721 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSCLLIQMPRFKGFPLFKKIFPSELEINITDL 780
Qy 774 LEDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTKNTQVHLHPKRLNKNPVSIL 833
Db 781 LEDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTKNTQVHLHPKRLNKNPVSIL 840
Qy 834 PKDLPDWDNRHGICPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNG 893
Db 841 PKDLPDWDNRHGICPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNG 900
Qy 894 FNIPOVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMVQSPMTSLYK 949
Db 901 FNIPOVTPCPEVGEYLKMSLEDLHLSDSRRIQGCARRLLCDAYMCMVQSPMTSLYK 956

RESULT 3

US-10-755-889-490

Sequence 490, Application US/10755889
Publication No. US20040171823A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
FILE OF INVENTION: PATHWAY
FILE REFERENCE: D0284 NP
CURRENT APPLICATION NUMBER: US/10/755,889
CURRENT FILING DATE: 2004-01-13
PRIOR APPLICATION NUMBER: U.S. 60/440,068
PRIOR FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: U.S. 60/469,757
PRIOR FILING DATE: 2003-05-12
NUMBER OF SEQ ID NOS: 823
SOFTWARE: Patent in version 3.2
SEQ ID NO 490
LENGTH: 956
TYPE: PRT
ORGANISM: Homo sapiens
US-10-755-889-490

Query Match 98.8%; Score 4971.5; DB 16; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQEKVTSPLYWEERIFVLLQECVTDKQTKLLKVPKSGSIGQYIDRSVGHRI 60
DB 1 MSSGLWSQEKVTSPLYWEERIFVLLQECVTDKQTKLLKVPKSGSIGQYIDRSVGHRI 60
QY 61 PSAGKKNQIGKILKEOPHNVFVDE-DVVEINEKFTTELLAITNCEERFSLFKNENRIS 119
DB 61 PSAGKKNQIGKILKEOPHNVFVDE-DVVEINEKFTTELLAITNCEERFSLFKNENRIS 120
QY 120 KGLQIDVGCVPVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEEGRGQGTGV 179
DB 121 KGLQIDVGCVPVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEEGRGQGTGV 180
QY 180 YQKQLFQCEDECQ-FVALDKLELIEDDDTALESYAGPDMQVLPPELPLEINSRVSLKG 238
DB 181 YQKQLFQCEDECQ-FVALDKLELIEDDDTALESYAGPDMQVLPPELPLEINSRVSLKV 240
QY 239 GTTIESGTIVFCVLPKESLGFVGVDMNDPIGNWDGRFDGV-LCSFACVESTILLHN 297
DB 241 GTTIESGTIVFCVLPKESLGFVGVDMNDPIGNWDGRFDGV-LCSFACVESTILLHN 300
QY 298 DIIP---ESVTQERRPPKLAFMSRGVGDGKSSSHNPKATGSTSDPGNR-RSELFTYLAG 353
DB 301 DIIPALSSVTQERRPPKLAFMSRGVGDGKSSSHNPKATGSTSDPGNRNRSSELFTYLAG 360
QY 354 SSVDSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLITENRPHSLP 413
DB 361 SSVDSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLITENRPHSLP 420
QY 414 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQSPPLAMPNGSHGLVGSVLAEVKE 473
DB 421 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQSPPLAMPNGSHGLVGSVLAEVKE 480
QY 474 NPPFGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFKVLKSCR 533
DB 481 NPPFGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFKVLKSCR 540
QY 534 PDSRFASLOPVNSQIERCNLSAFGGVLSVVEENTPPKMEKEGLETWIGKKGIQGHYNS 593
DB 541 PDSRFASLOPVNSQIERCNLSAFGGVLSVVEENTPPKMEKEGLETWIGKKGIQGHYNS 600
QY 594 CYLDDTLFCLFAFSSVLDTVLLRPKEKNDVEYYSQELLRTIEVNPLRIYGVVCAKIM 653
DB 601 CYLDDTLFCLFAFSSVLDTVLLRPKEKNDVEYYSQELLRTIEVNPLRIYGVVCAKIM 660
QY 654 KLRKILEKVEAASGFTSEBKDPEEFNLIFLHILRVEPLLKIRSAQKQVQDCYFYQIFME 713
DB 661 KLRKILEKVEAASGFTSEBKDPEEFNLIFLHILRVEPLLKIRSAQKQVQDCYFYQIFME 720

QY 714 KNEKVGVPITQOLLEWSFINSNLKFAEAPSCIIQMPRFQKDPKLPKKIIPPSLELNTDL 773
DB 721 KNEKVGVPITQOLLEWSFINSNLKFAEAPSCIIQMPRFQKDPKLPKKIIPPSLELNTDL 780
QY 774 LEDTPQRCRICGGLAMYECECYDDDPDISAGKIKQFCKTQNTQVHLHPKELNHNKYNPVS 833
DB 781 LEDTPQRCRICGGLAMYECECYDDDPDISAGKIKQFCKTQNTQVHLHPKELNHNKYNPVS 840
QY 834 PKDLPDWRHGCIPQNNMELFAVLCITSHYVAFVKYGHKDDSAWLFFDSMADRDGGONG 893
DB 841 PKDLPDWRHGCIPQNNMELFAVLCITSHYVAFVKYGHKDDSAWLFFDSMADRDGGONG 900
QY 894 FNIPOVTPCEVGEYKMSLEDLSLSDSRRIQCCARLLCDAYMCMYQSPMTSLYK 949
DB 901 FNIPOVTPCEVGEYKMSLEDLSLSDSRRIQCCARLLCDAYMCMYQSPMTSLYK 956

RESULT 4
US-10-921-707-9
Sequence 9, Application US/10921707
Publication No. US20050003447A1
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
APPLICANT: LAL, Preeti
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: BANDMAN, Olga
APPLICANT: CORLEY, Neil C.
APPLICANT: GUEGLER, Karl J.
APPLICANT: PATTERSON, Chandra
APPLICANT: AZIMZAI, Yalda
APPLICANT: BAUGHN, Mariah R.
TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
FILE REFERENCE: PF-0594 PCT
CURRENT APPLICATION NUMBER: US/10/921,707
CURRENT FILING DATE: 2004-08-19
PRIOR APPLICATION NUMBER: US/09/786,797
PRIOR FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: 09/156,470; unassigned: 60/131,321
PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 731
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 2363327
US-10-921-707-9

Query Match 76.7%; Score 3862; DB 16; Length 731;
Best Local Similarity 99.8%; Pred. No. 6.1e-315;
Matches 728; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 221 MOVLEPPELEINSRVSLKGGTIESGTIVFCVLPKESLGFVGVDMNDPIGNWDGRFDG 280
DB 1 MOVLEPPELEINSRVSLKGGTIESGTIVFCVLPKESLGFVGVDMNDPIGNWDGRFDG 60
QY 281 V-LCSFACVESTILLHNDIIIPESVTQERRPPKLAFMSRGVGDGSSSHNPKATGSTSD 339
DB 61 VQ-LCSFACVESTILLHNDIIIPESVTQERRPPKLAFMSRGVGDGSSSHNPKATGSTSD 120
QY 340 PGNR-RSELYFTYLANGSVDSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPP 398
DB 121 PGNRNRSSELFTYLANGSVDSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPP 180
QY 399 PVNSLTENRPHSLPESLTAMPNTNGSIGHSPLSLSAQSVMEELNAPVQSPPLAMPNG 458
DB 181 PVNSLTENRPHSLPESLTAMPNTNGSIGHSPLSLSAQSVMEELNAPVQSPPLAMPNG 240
QY 459 NSHGLEVGSIAEVKENPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFT 518

Abandoned - Failure to Resp.
to an OA.
still lost
foreign priority

```
Db 241 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACGCTDGTFRGRYFT 300
Qy 519 CALKALFVKLKS CRPDSRFSALQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLE 578
Db 301 CALKALFVKLKS CRPDSRFSALQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLE 360
Qy 579 IMIGKKGIQGHYNSCYLSTLFCPLPAPSSVLDTVLLRPKEKNDVYSETQELLATEIV 638
Db 361 IMIGKKGIQGHYNSCYLSTLFCPLPAPSSVLDTVLLRPKEKNDVYSETQELLATEIV 420
Qy 639 NPLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSA 698
Db 421 NPLRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSA 480
Qy 699 GQKVQDVCYQYFMENKXGVPTIQOLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL 758
Db 481 GQKVQDVCYQYFMENKXGVPTIQOLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL 540
Qy 759 FKIPFSLNITDLEDTPRCICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVH 818
Db 541 FKIPFSLNITDLEDTPRCICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVH 600
Qy 819 LHPKRLNHNKYNPVS LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAW 878
Db 601 LHPKRLNHNKYNPVS LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAW 660
Qy 879 LFPDSMADRDGGGNGFNIPQVTPCPVEGYLKMSLDLSRRITQGCARRLLCDAYMC 938
Db 661 LFPDSMADRDGGGNGFNIPQVTPCPVEGYLKMSLDLSRRITQGCARRLLCDAYMC 720
Qy 939 MYQSPTMSLYK 949
Db 721 MYQSPTMSLYK 731

RESULT 5
US-10-264-237-1609
; Sequence 1609, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1609
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (209)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (218)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (227)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (257)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1609
```

```
Query Match 25.3%; Score 1275; DB 15; Length 261;
Best Local Similarity 98.7%; Pred. No. 2.6e-98;
Matches 234; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 653 MKLRKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSA GQKVQDVCYQYFM 712
Db 1 MKLRKILEKVEAASGFTSEKDPESFLNLFPHILRVEPLLKIRSA GQKVQDVCYQYFM 60
Qy 713 EKNEKVGVTIQOLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL FKIPFSLNITD 772
Db 61 EKNEKVGVTIQOLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKL FKIPFSLNITD 120
Qy 773 LLEDTPRCICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVHLHPKRLNHNKYNPVS 832
Db 121 LLEDTPRCICGGLAMYECCYDDPDISAGKIKQFCKTCTNTQVHLHPKRLNHNKYNPVS 180
Qy 833 LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAW LFPDSMADRDG 889
Db 181 LKPLDPDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAW LFPDSMADRDG 237

RESULT 6
US-10-264-237-1610
; Sequence 1610, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1610
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (227)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1610

Query Match 23.3%; Score 1173; DB 15; Length 238;
Best Local Similarity 98.2%; Pred. No. 8.3e-90;
Matches 224; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 419 MPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPLPMPNGSHGLEVSLAEVKNPPFY 478
Db 1 MPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPLPMPNGSHGLEVSLAEVKNPPFY 60
Qy 479 GVIRWIGQPPGLNEVLAGELEDEACGCTDGTFRGRYFTCALKKALFVKLKS CRPDSRF 538
Db 61 GVIRWIGQPPGLNEVLAGELEDEACGCTDGTFRGRYFTCALKKALFVKLKS CRPDSRF 120
Qy 539 ASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMTICKKGIQGHYNSCYLDS 598
Db 121 ASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMTICKKGIQGHYNSCYLDS 180
Qy 599 TLFCLFAFSVLDTVLLRPKEKNDVEYSETQELLRTEIVNPLRIYGY 646
```


Db 181 TLFLCFXFSVLDVLRPKKNDVEYSETQELLRTVNPURINXY 228

RESULT 7

US-09-864-761-34675

Sequence 34675, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecomica-X-1

CURRENT APPLICATION NUMBER: US/09/864, 761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263, 6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 34675

LENGTH: 113

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC007728.1

OTHER INFORMATION: EXPRESSED IN BT47A, SIGNAL = 0.98

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4

OTHER INFORMATION: EST HUMAN HIT: A1130924.1, EVALUE 5.00e-02

OTHER INFORMATION: SWISSPROT HIT: Q03164, EVALUE 8.90e-01

US-09-864-761-34675

Query Match 11.9%; Score 597; DB 9; Length 113;

Best Local Similarity 100.0%; Pred. No. 6.1e-42;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 DFDSSPPLOPPVNSLTTRRPHSLPFLSTKMPNTNGSICHSPLSLSAQSVMEELNTAP 446

DB 1 DFDSSPPLOPPVNSLTTRRPHSLPFLSTKMPNTNGSICHSPLSLSAQSVMEELNTAP 60

QY 447 VQSPPLAMPNGSHGLEVGLAEVKNPPFYGVIRWIGOPPGNLVAGLEL 499

DB 61 VQSPPLAMPNGSHGLEVGLAEVKNPPFYGVIRWIGOPPGNLVAGLEL 113

RESULT 8

US-10-087-192-120

Sequence 120, Application US/10087192

Publication No. US20020182586A1

GENERAL INFORMATION:

APPLICANT: Morris, David W.

APPLICANT: Engelhard, Eric K.

TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

TITLE OF INVENTION: CANCER

FILE REFERENCE: 529452000122

CURRENT APPLICATION NUMBER: US/10/087,192

CURRENT FILING DATE: 2002-03-01

PRIOR APPLICATION NUMBER: US 09/747,377

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: US 09/798,586

PRIOR FILING DATE: 2001-03-02

NUMBER OF SEQ ID NOS: 2059

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 120

LENGTH: 354

TYPE: PRT

ORGANISM: Homo sapiens

US-10-087-192-120

Query Match 2.9%; Score 146; DB 13; Length 354;

Best Local Similarity 20.4%; Pred. No. 0.0031;

Matches 89; Conservative 48; Mismatches 147; Indels 152; Gaps 19;

QY 219 DTMQVELPPLPINSRVSLKG-----GETTESGTIVFCVDLPKRESLGYFVGVDMDN 269

DB 53 ETQEEFVDVDFRGERVWVNGKPCFIQFLGET-----QFAPGQ-----WAGIVLDE 98

QY 270 PIGNWGRFDGVLCSFACVESTILLHINDIIPESVTOERRPPPKLAFMSRGVGDGSSSHN 329

DB 99 PIGNWGRFDGVLCSFACVESTILLHINDIIPESVTOERRPPPKLAFMSRGVGDGSSSHN 329

QY 330 KPKATGSTDPGNRRSELFTLNGSSVDSPQSKSKNTWYIDEVAEDPAKSLTEISTDFD 389

DB 135 EDEANGLQITTFASRATSPCLCTSTASVMVSSPSTPSN-----IPQKPSQFA----- 179

QY 390 RSSPPLQPPVNSLTTRRPHSLPFLSTKMPNTNGSICHSPLSLSAQSVMEELNTAPVQE 449

DB 180 -AKEPSATPISNLT-----KTASSEISNLSAEGSIKK 211

QY 450 SPPLAMPNGSHGLEVGLAEVKNPPFYGVIRWIGOPPGNLVAGLELDEDEGAGCTDG 509

DB 212 -----GERELKIGDRVLVGGTGA--GVVRFLGETDFAKGEWCGVEL--DEPLGKNDG 259

QY 510 TFRGTRYFTCALKKALFVKLSCRPDSRFASLQVPSNQIERCNSLAFGYSILVEVEENTP 569

DB 260 AVAGTRYFQCPKYGLFA-----PVHKVTGIGF-----PSTTP 292

QY 570 PKMEKGL-EIMIGKKKGIQGHVNSCYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYYS 628

DB 293 AKANAVRRWMTTASLRKSPSASSLS-----MSSVASSVSSRPS----- 335

QY 629 TQELLRTETVNPRLRY 644

QY 450 SPPLAMPPGNHGLEVGS---LAEVKENPPFYGVIRWIGOPGLNEVLAGELEDEACGC 506
Db 417 -----EVGDQVLVAGQKQ-----GIVRFYKTDFAFGWYGIEL-DQPTGK 456
QY 507 TDGTRGTYFTCAKALFVKLSCRPDSR 537
Db 457 HDGSVFGVRYFTCPPRHGVEA-----PASR 481

RESULT 11

US-10-104-047-2990
; Sequence 2990, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2990
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2990

Query Match 2.7%; Score 135.5; DB 15; Length 306;
Best Local Similarity 22.7%; Pred. No. 0.019;
Matches 80; Conservative 30; Mismatches 90; Indels 153; Gaps 15;

QY 196 ALDKLEIEDDTALESDYAGPQDTMQVELPPIEINSRVSLKGETIESGTVFICDVLPG 255
Db 30 ALPKVTLPNYDNV-----PQNLMSAL-GLRLGDRVLLDQKT---GTLRFQGT-- 75
QY 256 KESLGFYGVDMNPIGNWDGPDGV---LCS-----FACVESTILLHINDIIPESVTOE 307
Db 76 EFASGQWGVDELDEPKGNDGSGVGVRYFCPPKQGLFASVK--ISKAVDAPPSSVTST 133
QY 308 RPPKLAFMRSRGVDKSGSSHNNKPKATGSTDPGNRRSELFTYLNSSVDSQPSKSKNT 367
Db 134 PTPRMDP--SRVTG-KGREHKGKTKTSPSLGSLQ----- 169
QY 368 WYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTENRPHSLPFLTKMNTNGSIG 427
Db 170 -----RDGAKA----- 175
QY 428 HSPLSLSAQSVMEELNAPVQESPPLAMPNGNHHGLEVGS---LAEVKENPPFYGVIRWI 484
Db 176 -----EVGDQVLVAGQKQ-----GIVRFY 194
QY 485 GQPPGLNEVLAGELEDEACGCTDGTFRGTRYFTCAKALFVKLSCRPDSR 537
Db 195 GKTDFAPGWYGIEL-DQPTGKHDSGVFGVRYFTCPPRHGVEA-----PASR 240

RESULT 12

US-09-764-875-857
; Sequence 857, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 857
; LENGTH: 721

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (358)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (485)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-875-857

Query Match 2.7%; Score 135.5; DB 11; Length 721;
Best Local Similarity 23.1%; Pred. No. 0.076;
Matches 82; Conservative 44; Mismatches 132; Indels 97; Gaps 14;
QY 193 GFVALDKLEIEDDD-----TALE-----SDYAGPQDTMOVELPPL 228
Db 132 GYEDLDTFKLEEDDELNIRDPHEHRAVLLTAVELLQYDSDSDQSGQEKLLVDQGL 191
QY 229 E-INRSVS--LKGETIESGTVFICDVLPGKESLGYFVGVDMNDPIGNWDGFRDGVLCSE 285
Db 192 SGCSPRDSGCYSESENLENGKTRKASLLSAKSTEPSLKSFSRNQNGY----- 240
QY 286 ACVESTILLHINDIIPESVTQERRPPKLAFMRSRGVDKSGSSHNNKPKATGSTDPGNRRS 345
Db 241 ---PTPLMKSGDALKQGEGR-----LGGGLAPDTSKCDPPGVTLNKK--NRRS 287
QY 346 -----ELFYTLNG-SSVDSQPSKSKNTWIID-----EVAEDPAKSLTEISTDFD 389
Db 288 LPVSIICRSCTLEPGQVDTWPRSHSLDLDLQVEFGAEQDVPTETVTEPPQIVPEVPOKTT 347
QY 390 RSSPPLQPPVNS-----LTTENRPHSLPFLTKMNTNGSICHSPILSLSAQSVMEELN 443
Db 348 ASSTKAQPLEXDSAVDNALLTQSKRPFSEPKLT-TKKLEGSIAASRGGL----- 396
QY 444 TAPVQESPPL-----AMPNGNHHGLEVGSLEAEVKENPPFYGVIRWIGOPPLN 491
Db 397 -----SPQCLPRNYDAQPPGAKHGLARTPLGHRKKGHEFGTHHPLGTKEGYD 445

RESULT 13

US-10-282-122A-66115
; Sequence 66115, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347

```
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66115
; LENGTH: 2273
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-282-122A-66115

Query Match      2.7%; Score 135.5; DB 15; Length 2273;
Best Local Similarity 20.4%; Pred. No. 0.5;
Matches 177; Conservative 112; Mismatches 315; Indels 265; Gaps 42;

QY 34 QTOKLKVPGSIGQVQIDRSVGHSRIPSAKGNKQIGLKILEQPHAVLFDVDEVDVEINE 93
Db 1345 QTGSTISSPQDVG-----ISSKISIDRAQNRYSQESKQVYEQKGVTVVAISVPVVTVM 1399
QY 94 KFTPELLAIINC-----EERFSLPKNRLSKGLQIDV-----GCPVKVQLRS 136
Db 1400 GAVDAKAVQTVGKSKNSRVNMAAANLNKGVDSVALYNAARNPKKAAGQGISVSVTY 1459
QY 137 GEEKFPVVRFRGPLLAERTVSGIFFG--VELLEEGRGQGT-----DGYVQCK----- 183
Db 1460 GEQKNTSESIKQTVQEGKITG--GGKVSLTASGAKDSRITITGSDVYGGKGTSLKA 1516
QY 184 -----QLFQCBDCGF-----VALDKLELIEDDDTALESYA---GPGDTMQVE 224
Db 1517 ENAVQIEAARQTHQERSENKSAAGNAGVAIAINKGISFGFTA--GANYGKGYNGDETAYR 1575
QY 225 LPPLNLSRVSLKGET--IESG--TWIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV 281
Db 1576 -----NSHTGSKDSQTAIESGDDTVIKGQLKGG-----VGVTAES----- 1612
QY 282 LCSFACVESTILHINDIIPESVTOERRPPKLPFMSRGVGDGSSSHNKPATGSTSDPG 341
Db 1613 -----LHIESLQDTAFVKQENVSQVTVGYGFSVGGSVNRK-----SSSDYA 1657
QY 342 --NRRSELF-----YTIN-----GSSVDSQPQSKSN-----TWYID-EVAEDPAK 379
Db 1658 SVNEQSGIFAGGQGYRVRNGKTGLVGAUVSD--ADKSKNLLKTSEIWHKDIQNHASAAA 1716
QY 380 SLTEISTDFRSPPPLOPPVNSLTNRRPHSLPFSLTKMPNTNGSIGHSPLSL----- 433
Db 1717 SALGLSGGFSYSPK-----TSQOYS-----TKAEABIGTIGKRPVSLMRFDQV 1760
QY 434 SAQSVMEELNATPVQES-----PPLAMPNGSHGLEVG-----466
Db 1761 SAKD--DELMEKYSRERIEKETFKKANLNQNNAGGLKFLKQNDIHSNDKYALAKMGLG 1818
QY 467 -SLAEVKENPPFYGVIRWIOQPPCLNE---VLAGELEDECACTDGTGTRGTFTCALK 522
Db 1819 NLJGNAKESSESRSQITRSV-----ISEGDWQIASAQGRKNIAIEKGTSSAHLAKAKAD 1873
QY 523 KALFVKLKSRCPSDRFASLPQVSNQIERNCSLAFGGYLVSEV-----VEEN 567
Db 1874 EGLLKEVELNRDVAK-----EFINETLIGGIADAVRSQFTAEHRLMTFKMDEN 1922
QY 568 TPKMEKEGLEIMTKKKGIGQHNSCYLDSTLFCLPFSSVLD-----TWLLRPK 618
Db 1923 GER-IDKQLEEDINKQ-----FNSVKLKEKFPASFKDYNWAYKAIGNIYELR 1970
QY 619 EKNDVEYSETQELLRTIEVNLRIYGVVCATKIMKLKILEKVEAASGFTSEK-----D 674
```

```
Db 1971 EVSD-----QERNLKTARTYDDET--GKTVEKIVGVNGVFNFIQAARFAAQYVGRFN 2024
QY 675 PBE-----FLNLFHHLRVEPLLKIRSAGQVQDCYFYQIFMEKNEKVGVTIQOLLE 728
Db 2025 PEKNRYERYENVYFLH---NPETNTRG-----FSKLPELAAVAFAFHOMLE 2066
QY 729 WSPFINSNLKPAEAPSCLIQ--MPRFGKD 755
Db 2067 GAKIGNKTVIGLSNGLALGNIMEDYKGD 2095

RESULT 14
US-09-764-868-731
; Sequence 731, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 731
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-731

Query Match      2.7%; Score 134.5; DB 9; Length 721;
Best Local Similarity 23.1%; Pred. No. 0.092;
Matches 82; Conservative 44; Mismatches 132; Indels 97; Gaps 14;

QY 193 GFVALDKLELIEDD-----TALE-----SDVAGPGDTMQVELPPL 228
Db 132 GYEDLDTFKLEBEDDELNIRDPEHRAVLTLTAVELLQEYDSDNSDQSGOEKLLVDSSQL 191
QY 229 E-INSRVS--LKGETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGVLCSE 285
Db 192 SGCSPRDSGCYSESENLENGKTKASILLSAKSSTESLSKFSRNQLGN-----240
QY 286 ACVESTILHINDIIPESVTOERRPPKLPFMSRGVGDGSSSHNKPATGSTSDPGNRRS 345
Db 241 ---PTLPLMKSGDALKQGBEGR-----LGGGLAPDTSKCDPPGVGTGLNK---NRRS 287
QY 346 -----ELFYLNG--SSVDSQPQSKSNWTYID-----EVAEDPAKSLTEISTDFD 389
Db 288 LPVSIICRSCETLEGPTQVDTWPRSHSLDDLLQVEFGAEQDVPTTEPPQIVPEVPOKTT 347
QY 390 RSSPPQPQPPVNS-----LTTENRPHSLPFSLTKMPNTNGSIGHSPLSLSAQSVMEELN 443
Db 348 ASSTKAQPLEQDSAVDNALLLTQSKRFPSEPKLT--TKLEGSTAASGRGL-----396
QY 444 TAPVQESPPL-----AMPPGNHGLEVGSLEAVKENPPFYGVIRWIGPPGLN 491
Db 397 -----SPQCLPRNYDAQPPGAKGHLARTPLEGHRKGHEFEGTHPLGTKEGVD 445

RESULT 15
US-09-764-875-1140
; Sequence 1140, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: FJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
```

```

; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1140
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-875-1140

Query Match      2.7%; Score 134.5; DB 11; Length 721;
Best Local Similarity 23.1%; Pred. No. 0.092;
Matches 82; Conservative 44; Mismatches 132; Indels 97; Gaps 14;

QY 193 GFVALDKLEIEDDD-----TALE-----SDVAGGDTMQVELPPL 228
Db 132 GYEDLDTFKLEEDLDLNLNIRDPHEHRAVLLTAVELLQYDSDNSDQSGQEKLVDSQGL 191
QY 229 E-INSRVS--LKGETIESGTWIFCDVLPKESLGYFVGVDMDNPIGNMDGRFDGVLCSP 285
Db 192 SGCSPRDSCYESSLENLENGKTRKASLLSAKSTEPSLKSFSRNQLGNY-----240
QY 286 ACVESTILHINDIIPESVTQERRPPKLAFMRSRGVGDGSGSSHNKPKATGSTDGPNRRS 345
Db 241 ---PTLPLMKSGDALKQGEGR-----LGGGLAPDTSKCDPPGVTLNKK---NRRS 287
QY 346 -----ELFYTLNG--SSVDSQPSQSKNTWYID-----EVAEDPAKSLTEISTDFD 389
Db 288 LPVSIICRSCETLEGPTQVDTWPRSHSLDDLQVEPGAEDVPTETPPQIVPEVPQKT 347
QY 390 RSSPPLQPPPVNS-----LTTENRRFHSLPFSLTKMPTNGSIGHSPLSLSAQSVMEELN 443
Db 348 ASSTKAQPLEQDSAVDNALLTQSKFSEFQKLT-TKKLEGSIAASGRGL-----396
QY 444 TAPVQESPPPL-----AMPPEGNSHGLEVGSLAEVKENPPFYGVIRWIGQPPGLN 491
Db 397 -----SPPQCLPRNYDAQPPGAKHGLARTPLEGHRKHGHEFGTHHPLGTKEGVD 445

```

Search completed: April 18, 2005, 15:35:35
Job time : 100 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 18:28:51 ; Search time 1146 Seconds
(without alignments)
5023.830 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPWEERI.....RLLCDAYMCYQSPWMSLYK 949

Scoring table:

BLOSUM62	Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5	
Fgapop 6.0 ,	Fgapext 7.0	
Delop 6.0 ,	Delext 7.0	

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09671687/runat_18042005_115020_17247/app_query.fasta_1.1095
-DB=Published Applications NA -QWMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODES=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USRR=US09671687 @CGN 1 1 723 @runat_18042005_115020_17247
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4971.5	98.8	5371	10	US-09-851-673-3	Sequence 3, Appli
2	4971.5	98.8	5371	18	US-10-755-889-489	Sequence 25, Appl
3	4942	98.2	4527	18	US-10-921-707-25	Sequence 230, App
4	3676	73.0	2523	15	US-10-037-270-290	Sequence 230, App
5	3676	73.0	2523	17	US-10-117-722-290	Sequence 3, Appli
6	2468	49.0	2116	18	US-10-761-370-3	Sequence 204, App
7	1778	35.3	1151	17	US-10-264-237-204	Sequence 205, App
8	1494	29.7	1013	17	US-10-264-237-205	Sequence 205, App
9	1102	21.9	617	11	US-09-969-034-4253	Sequence 4253, Ap
10	879	17.5	557	11	US-09-969-034-4254	Sequence 4254, Ap
11	826	12.4	376	9	US-09-983-965-1633	Sequence 1633, Ap
12	545	10.8	309	9	US-09-864-761-18145	Sequence 18145, A
13	512	10.2	483	9	US-09-864-761-1386	Sequence 1386, Ap
14	412	8.2	425	10	US-09-918-995-35878	Sequence 35878, A
15	160.5	3.2	4847	17	US-10-152-319A-1764	Sequence 1764, Ap
16	151	3.0	5857	10	US-09-873-367C-305	Sequence 305, App
17	151	3.0	5857	19	US-10-843-641A-305	Sequence 305, App
18	148	2.9	3051	17	US-10-371-905A-3	Sequence 3, Appli
19	146	2.9	1380	13	US-10-087-192-119	Sequence 119, App
20	146	2.9	5563	18	US-10-723-860-1545	Sequence 1545, App
21	146	2.9	5607	18	US-10-723-860-5973	Sequence 5973, Ap
22	143	2.8	3287	17	US-10-275-595A-47	Sequence 47, Appl
23	143	2.8	3370	17	US-10-108-260A-308	Sequence 308, App
24	143	2.8	3592	10	US-09-814-353-20078	Sequence 20078, A
25	138	2.7	2154	17	US-10-369-493-45974	Sequence 45974, A
26	138	2.7	2451	18	US-10-437-963-30316	Sequence 30316, A
27	136.5	2.7	3409	18	US-10-723-860-6166	Sequence 6166, Ap
28	135.5	2.7	2618	17	US-10-104-047-1020	Sequence 1020, App
29	135.5	2.7	3103	11	US-09-764-875-259	Sequence 259, App
30	135.5	2.7	6822	17	US-10-282-122A-29931	Sequence 29931, A
31	135	2.7	2643	9	US-09-350-874-56	Sequence 56, Appl
32	135	2.7	2643	15	US-10-106-989-56	Sequence 118, App
33	134.5	2.7	3103	9	US-09-764-868-118	Sequence 542, App
34	134.5	2.7	3103	15	US-09-764-875-542	Sequence 1329, Ap
35	134	2.7	1099	15	US-10-106-698-1329	Sequence 919, App
36	134	2.7	2855	17	US-10-094-749-919	Sequence 919, App
37	133	2.6	3245	17	US-10-120-988-352	Sequence 352, App
38	132.5	2.6	1416	13	US-10-087-192-116	Sequence 116, App
39	132	2.6	3338	19	US-10-491-213-65	Sequence 65, Appl
40	130	2.6	2676	17	US-10-276-774-450	Sequence 450, App
41	129.5	2.6	4839	18	US-10-363-829-121	Sequence 121, App
42	129.5	2.6	10531	14	US-10-160-758-10	Sequence 10, Appl
43	129.5	2.6	10759	17	US-10-210-172-49	Sequence 49, Appl
44	128.5	2.6	2446	9	US-09-801-275-1	Sequence 1, Appli
45	128.5	2.6	2446	16	US-10-170-789-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-09-851-673-3
; Sequence 3, Application US/09851673
; Publication No. US20030165985A1
; GENERAL INFORMATION:
; APPLICANT: Derry, Jonathan
; APPLICANT: Fanslow, William
; APPLICANT: Dougall, William
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
; FILE REFERENCE: 3198
; CURRENT APPLICATION NUMBER: US/09/851.673
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5371
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

LOCATION: (392)..(3262)
US-09-851-673-3

Alignment Scores:

Pred. No.: 0 Length: 5371
Score: 4971.50 Matches: 948
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 98.76% Indels: 7
DB: 10 Gaps: 5

US-09-671-687A-3 (1-949) x US-09-851-673-3 (1-5371)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 392 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAGTCACTTCCCTTACTTGGGAAGCGGATT 451
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 452 TTTTACTTGCTTCTTCAAGATGCGGTTACAGACATATATCAAGATCGTTCGTGGGCAATCAAGGATT 571
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 512 GTACCGAAGGAGATGACAGATATATCAAGATCGTTCGTGGGCAATCAAGGATT 571
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnGlnProHisAla 80
DB 572 CTTCTGCAAGGCAAGAAATAGATGGATTGATTAATAATCTAGAGCAACCTCATGCA 631
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeu 99
DB 632 GTTCTCTTTGTTGATCAAAAGGATGTTGTAGAGATAATGAAAGTTTCACAGAGTTTACTT 691
QY 100 LeuAlaIleThrAsnGlyGlnGluArgPheSerLeuPheLysAsnArgLeuSer 119
DB 692 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCAGCTCTTTTAAAAACAGAAACAGACTAAGT 751
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 752 AAAGGCTTCANATAGACGTGGGCTGTCTCTGTAAGATGACGTGAGATCTGGGGAGAA 811
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 812 AAATTTCTGGAGTTGTACGTTTCAGAGACCCCTGTAGCAGAGAGGACAGTCTCCGA 871
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
DB 872 ATATTTCTTTGGAGTTGAAATTTGCTGGAAGAAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 931
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 932 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGATTTGGAC 991
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 992 AAGCTAGAACTCATAGAGATGATGACCTGCAATTTGGAAGTGTATACGAGGTCCTGGG 1051
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1052 GACACAATGCGAGTCGAATCTCTCTCTTGGAAATAAATCCAGAGTTTCTTTGAAAGTT 1111
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1112 GGAAGAAACAATAGAAATCTGGAACAGATTATATCTGTGATGTTTGGCAGGAAAGAAGC 1171
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1172 TTAGGATATTTTGTGTGGACATGATTAACCTTATTTGGCACTGGGATGGAAGATT 1231
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1232 GATGGAGTGCAGCTTTGTAGTTTGGTGTGTTGAAAGTACAAATCTTATTGCATCAAT 1291

QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
DB 1292 GATATCATCCAGCTTTATCAGAGAGTGTGACCCAGGAAAGGAGGCTTCCCAACTTGGC 1351
QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
DB 1352 TTTATGTCAAGAGGTGTTGGGCAAGGTTTCCAGTCAATTAATTAACCAAGGCTTACA 1411
QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
DB 1412 GGATCTACCTCAGACCTCGGAATAGAACAGATCTGAATATTTTATATACCTTAAATGG 1471
QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
DB 1472 TCTTCTGTTGACTCACCAACATCCAAATCAAAATAACATGATGATGATGAAGTT 1531
QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
DB 1532 GCAGAAAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCACCA 1591
QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
DB 1592 CCACTCCAGCTCTCTCTGTGAACCTCACTGACCACCGAGAACAGATTCACACTCTTTTACA 1651
QY 414 PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
DB 1652 TTCAGTCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGCTCCACTTCTCTG 1711
QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGlnSerProProLeu 453
DB 1712 TCAGCCAGCTGTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCCCTG 1771
QY 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
DB 1772 GCATCCCTCTCGGAACTCACATGTTGTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAG 1831
QY 474 AsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
DB 1832 AACCTCTCTTCTATGGGTAAATCCGTTGGATCCGTCAGCCACCCAGGACTGAATGAAGTG 1891
QY 494 LeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
DB 1892 CTGCTGAGCTGGAATGGAAGATGATGTGCGAGGCTGTACGGATGGAACCTTCAGAGGC 1951
QY 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
DB 1952 ACTCGGTATTTTCACTGTGCCCTGGAAGAGCGCTGTTGTGAAACTGGAAGAGCTGCAGG 2011
QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
DB 2012 CTTGACTTAGGTTTGCACTTTCAGCCCGGTTTCCANTCAGATTGAGCGCTGTAACTCT 2071
QY 554 LeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
DB 2072 TTAGCATTTGGAGGCTTACTTAAGTAGAGTAGTAGAGAAATACTCCACCAAAATAATGAA 2131
QY 574 LysGluGlyLeuGluLeuMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer 593
DB 2132 AAAGAAGGCTTGGAGATAATGATTGGGAAGAAGAGCATCCAGGGTCAATTAACAATCT 2191
QY 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
DB 2192 TGTACTTAGACTCAACCTTATTTCTGCTTATTTTGTCTTATTTTGTCTGTGGACACTGTG 2251
QY 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu 633
DB 2252 TTACTTAGACCCAAAGAAAGACGATGATAGATATTTATAGTGAACCCCAAGAGCTACTG 2311
QY 634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
DB 2312 AGGACAGAAATTTTAATCTCTGAGATATATGATATGTGTGTGTCACAAAAAATATG 2371
QY 654 LysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLys 673

Db 1172 TTAGGATATTTTCTTGGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAAGATTT 1231
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1232 GATGGAGTGCAGCTTGTAGTTTGGCTGTGTGAAAGTACAATTCTATTGGACATCAAT 1291
Qy 298 AsnIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
Db 1292 GATATCATCCCAAGCTTTATCAGAGAGTGTGACGACAGAAAGGAGGCTCCCAAACTTGCC 1351
Qy 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
Db 1352 TTTATGTCAAGAGGTGTGGGACAAAGTTTCATCCAGTCATATAAACAAGAGGTACA 1411
Qy 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
Db 1412 GGATCTACCTCAGACCCCTGGAATAAGAACAGATCTGAATTATTTATACCTTAATATGG 1471
Qy 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
Db 1472 TCTTCTGTGTGACTCACACCAATCCAAATCAAAATAATACATGGTACATTTGATGAAGTT 1531
Qy 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
Db 1532 GCAGAAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCCGTTCTTCACCA 1591
Qy 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
Db 1592 CCACCTCCAGCCTCTCTGTGAATCCTGACCTCAGTACCCAGACAGATCCACTCTTTACCA 1651
Qy 414 PheSerLeuThrLysMetProLeuThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
Db 1652 TTCAGTCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTCG 1711
Qy 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu 453
Db 1712 TCAGCCAGTGTGTAAATGGAAGAGCTAAACACTGCACCCGCTCCAGAGAGTCCACCCCTTG 1771
Qy 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1772 GCCATGCCCTCTGGGAATCTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAG 1831
Qy 474 AsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
Db 1832 AACCTCTCTTATGGGTAAATCCGTGTGATCGGTGAGCCACGACGACTGAATGAAGTG 1891
Qy 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db 1892 CTCGCTGGAGCTGGAAGATGAGTGTGACGGCTGTACGGATGGAACCTTCAGAGGC 1951
Qy 514 ThrArgTyrPheThrCysAlaLeuLysAlaLeuPheValLysLeuLysSerCysArg 533
Db 1952 ACTCGGTATTTCACTGTGCCCTTGAAGAGGCGCTGTTGTGAACCTGAAGAGCTGAGG 2011
Qy 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 2012 CCTGACTCTAGGTTTCATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTGAATCT 2071
Qy 554 LeuAlaPheGlyTyrLeuSerGluValValGluAsnThrProProLysMetGlu 573
Db 2072 TTAGCATTTGGAGGCTACTTAAGTAGTAGTAGTAAGAAATAACTCCACCAAAATGGAA 2131
Qy 574 LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer 593
Db 2132 AAAGAAGGCTTGGAGATAATGATTGGGAAGAAAGAGGCAATCCAGGCTCAATTACAATCT 2191
Qy 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2192 TGTACTTAGACTCAACCTTATTCTGCTTATTGCTTTTAGTTCTGTTGGACACTGTG 2251
Qy 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 633
Db 2252 TTACTTTAGACCCCAAGAAAGACGATGTAGATAATTATAGTGAACCCCAAGAGCTACTG 2311

Qy 634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
Db 2312 AGCACAGAAATTTGTTAATCTCTCAGATATATGATGATGTCGCCACAAAAATATG 2371
Qy 654 LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
Db 2372 AAACCTGAGAAATACTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAA 2431
Qy 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2432 GATCTGAGGAATTTCTTGAATATTCTGTTTCATCATATTATTAAAGGTAGAACCTTTGCTA 2491
Qy 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2492 AAAATTAAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGGAA 2551
Qy 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn 733
Db 2552 AAAAATGAGAAAGTTGGCGTTCACCAATTCACAGCTTGTAGAATGGTCTTTTATCAAC 2611
Qy 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
Db 2612 AGTAACCTGAATTTTCAGAGGACCATCATGTCATATTTCAGATGCCCTCGATTGGA 2671
Qy 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2672 AAAGACTTTAAACTATTAAAAAAATTTTCTCTCTGGAATTAATAACAGATTTA 2731
Qy 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2732 CTTGAAGACACTCCACAGAGTCCCGGATATGTGGAGGGCTTGCAATGTATGAGTGTAGA 2791
Qy 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2792 GAATGCTAGACATCCGACATCTCAGCTGGAAAAAATCAAGCAGTTTGTGAAACCTGC 2851
Qy 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2852 AACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATAACCCAGTGTCACTT 2911
Qy 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2912 CCCAAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTCCAGAAATATGGAG 2971
Qy 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2972 TTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTGTGAAGTATGGGAAG 3031
Qy 874 AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 3032 GAGCATTTCTGCTGGCTCTTCTTTGACAGCATGCGCGATCGGATGGTGGTCAGAAATGGC 3091
Qy 894 PheAsnIleProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeu 913
Db 3092 TTTCAACATTTCTCAAGTCAACCCATGCCAAGAGTAGAGAGTACTTTGAAGATGCTCTG 3151
Qy 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCys 933
Db 3152 GAAGACCTGCTATTTCTTGGACTCCAGGAGAAATCCAAGGCTGTGCACCAAGACTGCTTTGT 3211
Qy 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3212 GATGCATATATGTCATGTACAGAGTCCAAACATGAGTTGTGTACAAA 3259

RESULT 3

US-10-921-707-25
; Sequence 25, Application US/10921707
; Publication No. US20050003447A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom

APPLICANT: YUE, Henry
 APPLICANT: HILLMAN, Jennifer L.
 APPLICANT: BANDMAN, Olga
 APPLICANT: CORLEY, Neil C.
 APPLICANT: GUEGLER, Karl J.
 APPLICANT: PATTERSON, Chandra
 APPLICANT: AZIMZAI, Yalda
 APPLICANT: BAUGHN, Mariah R.
 TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
 FILE REFERENCE: PF-0594 PCT
 CURRENT APPLICATION NUMBER: US/10/921,707
 CURRENT FILING DATE: 2004-08-19
 PRIOR APPLICATION NUMBER: US/09/786,797
 PRIOR FILING DATE: 2001-06-25
 PRIOR APPLICATION NUMBER: 09/156,470, unassigned; 60/131,321
 PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
 NUMBER OF SEQ ID NOS: 32
 SOFTWARE: PERL Program
 SEQ ID NO 25
 LENGTH: 4527
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No: 2363327
 US-10-921-707-25

Alignment Scores:
 Pred. No.: 0 Length: 4527
 Score: 4942.00 Matches: 944
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.17% Indels: 5
 DB: 18 Gaps: 4

US-09-671-687A-3 (1-949) x US-10-921-707-25 (1-4527)

QY 4 GlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIlePheTyrLeu 23
 DB 270 GGCCTATGAGCCAGAAAAAGTCACTTCAACCTCTGGAAGAGCGGATTTTACTTG 329
 QY 24 LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLysValProLys 43
 DB 330 CTTCTTCAAGATGCGACGCTTACAGACAAACAAACAAAGCTCTTAAAGTACCGAA- 388
 QY 44 GlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIleProSerAla 63
 DB 389 GGAAGTATAGGACAGTATATTCAGATCGTTCTGTGGGGCAITTCAGGATTCCTCTGCA 448
 QY 64 LysGlyLysLysAsnGlnIleGlyLysLysIleLeuGluGlnProHisAlaValLeuPhe 83
 DB 449 AAAGGCAAGAAAAATCAGATTGGATTAAAAATTCAGAGCAACTCATGCACTTCTTT 508
 QY 84 ValAspGlu---AspValValGluLysLeuAsnGluLysPheThrGluLeuLeuAlaIle 102
 DB 509 GTTGATGAAAGGATCTGTAGAGATAAATGAAAGTTACAGAGTTACTTTTGGCAATT 568
 QY 103 ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeu 122
 DB 569 ACCAATGTGAGAGAGGTTGACCCCTGTTTAAAAACAGAAACAGACTAAGTAAAGGCCCTC 628
 QY 123 GlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGluLysPhePro 142
 DB 629 CAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAAAAATTTCTCT 688
 QY 143 GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe 162
 DB 689 GGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGTCTCCGGAATATCTTT 748
 QY 163 GlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGly 182
 DB 749 GGAGTTGAATGTGGAAGAGGCTCGTGTCAAGGTTTCTACCTGACGGGGTGTACCAAGGG 808

QY 183 LysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAspLysLeuGlu 201
 DB 809 AAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGATTGGAACAAGCTAGAA 868
 QY 202 LeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet 221
 DB 869 CTCATAGACATGATGACACTGCATTGGAAGTGAATTACGCGAGGTCTCTGGGACACAATG 928
 QY 222 GlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGlyGlyGluThr 241
 DB 929 CAGGTGGAATCTCTCTCTTGGAAATAAATCCACAGAGTTCTTTCGAAGTTTGGAGAAACA 988
 QY 242 IleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyr 261
 DB 989 ATAGAATCTGGAACAGTTATTTCTGTGATGTTTTGCCAGAAAGAAAGCTTAGGATAT 1048
 QY 262 PheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPheAspGlyVal 281
 DB 1049 TTTGTTGGTGTGACATGGATTAACCTATTGGCACTGGGATGGAAGATTGATGGAGTG 1108
 QY 282 ---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIle 300
 DB 1109 CAGCTTTGTAGTTTGGTGTGTTGAAAGTACAATTCTATTGCACATCAATGATATCATC 1168
 QY 301 ProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyVal 320
 DB 1169 CCAGAGAGTGTGACGAGGAAAGAGGCCTCCCAAACTTGCCTTTATGTCGAAGAGGTGT 1228
 QY 321 GlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspPro 340
 DB 1229 GGGGACAAAGGTTTCACTCCAGTCATAATAAACCAAGGCTACAGGATCTTACCTCAGACCCT 1288
 QY 341 GlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGln 359
 DB 1289 GGAATAGAAACAGATCTGAATTTTATACCTTAAATGGGTCTTCTGTGACTCACA 1348
 QY 360 ProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLys 379
 DB 1349 CCACAATCCAAATCAAAAAATACATGATGATGAAGTTGCAAGAGACCCCTGCAAAA 1408
 QY 380 SerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGlnProProPro 399
 DB 1409 TCTCTTACAGAGATATCTACAGACTTTGACCGTTCCTTCCACCACCTCCACCTCTCTCT 1468
 QY 400 ValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMet 419
 DB 1469 GTGAATCTCACTGACCCAGAGACAGATTCCACTCTTTACCATTCAGTCTCACCAGATG 1528
 QY 420 ProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMet 439
 DB 1529 CCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAGCTCTGTAATG 1588
 QY 440 GluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsn 459
 DB 1589 GAAGAGCTAAACACTGACCCGTCACAGAGAGTCCACCTTTGGCCATGCTCTCTGGGAAC 1648
 QY 460 SerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPheTyrGly 479
 DB 1649 TCATCTGTCTAGAGTGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTTCTATGGG 1708
 QY 480 ValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeu 499
 DB 1709 GTAATCCGTTGGATCGGTGACCCACAGGACTGAATGAAGTCTCGTGGACTGGAACTG 1768
 QY 500 GluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCys 519
 DB 1769 GAAGATGAGTGTGAGGCTGTACGATGGAACCTTTCAGAGGCACTCGGTATTTTCACTGT 1828
 QY 520 AlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAla 539
 DB 1829 GCCTTGAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGACGCTGACTCTAGGTTTGA 1888
 QY 540 SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyTyr 559

← SER 25 = here but this =
 371 filed
 in 1999
 not
 here
 SER ID 21
 in this
 provisional.

QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333
DB 225 GCTTTATGTCAAGAGGTGTGGGCAAGGTTCTCCAGTCATATAAACAAGGCT 284
QY 334 ThrGlySerThrSerAspProGlyAsnArg- - - ArgSerGluLeuPheTyrThrLeuAsn 352
DB 285 ACAGGATCTACCTCAGACCTCGAATAGAACAGATCTGAATATTTATACCTTAAT 344
QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu 372
DB 345 GGGTCTCTCTGTTGACTCACAAACACAAATCCAAATCAAAAATACATGGTACATTGATGA 404
QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392
DB 405 GTTGCAAGAGACCTCCAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTCTCA 464
QY 393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412
DB 465 CCACCACTCCAGCTCTCTGTGAATCTACTGACCAACCGAGAACAGATTCACACTCTTTA 524
QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
DB 525 CCATTGAGTCTCACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCT 584
QY 433 LeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProPro 452
DB 585 CTGTGAGCCAGCTCTGTAATGGAAGAGCTAAACACTGCACCCCTCCAAAGAGTCCACCC 644
QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
DB 645 TTGGCCATCCCTCTGGGAACCTCACATGGTCTGAGAAGTGGGCTCATTTGGCTGAAGTTAAG 704
QY 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
DB 705 GAGAACCTCTCTTATGGGTAAATCCGTGGATCGGTGAGTCCAGCACCAAGGACTGAATGA 764
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTGCTCGCTGGACTGGAATGGAAGATGAGTGTGCAGGCTGTACCGATGGAACCTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
DB 825 GGCACCTCGGTATTTCACTGTGCCCTGGAAGAGCGCTGTTGTGAACCTGAAGAGCTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGCCCTGACTTAGGTTTCATTCATTTGCAGCCCGTTTCCNATCAGATTGAGCCCTGTAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTGGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGGCTTGAGATAATGATTTGGGAGAGAAAGAGGATCCAGGGTCAATTACAAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTATTACTAGACTCAACCTTATTCTGCTTATTTGCTTTTATGTTCTGTCTGGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTTACTTAGACCAAGAAAGAACGATGTAGATATTTATAGTGAACCCCAAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACAGAAATGTTAATCTCTGAGATATATGATATGTGTGTGTCACAAAAAT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAAACTGAGGAAAAATCTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGA 1304

QY 673 LysAspProGluGluPheLeuAenIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCCTGAGGAATCTTTGAATATTTCTGTTTTCATCATATTTTAAGGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAATAATAGATCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCTTCCCACAATTCAGCAGTTGTTAGAAATGCTCTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTGAGAGGCACCATCATGTCTGATTTTTCAGATGCTCGATTT 1544
QY 753 GlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAAACTATTTTAAAAAATTTTCTCTCTGGAATTTAAATATAACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTTGAAGACACTCCAGACAGTGCAGATATGTGAGGGCTTGCATATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAAGATGCTTACAGCAGTCCGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAACCC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAGTCCACCTTCATCCAGAGGCTGAATCATATAATATAACCCAGTGTC 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGACTTACCCGACTGGACTGGAGACACGCTGCATCCCTTGCAGAAATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTATTTCGTCTCTCATAGAAACAAAGCCACTATGTTGCTTTTGTGAAGTATGGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGAGCAATTCGCTGGCTCTCTTTTGACAGCATGCGGATCGGATGGTGGTCAGAAAT 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
DB 1965 GGCTTCAACATCTCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeu 932
DB 2025 CTGGAAGACCTGCATTCCTTGGACTCCAGAGAAATCCAAAGGCTGTCACGAAAGACTGCTT 2084
QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 2085 TGTGATGCATATATGTGCATGTACCAGAGTCCCAATAGATGTTGTACAA 2135

RESULT 5

; Sequence 290, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Poly peptides
; FILE REFERENCE: 784CJP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19

```
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 290_2523
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2138)
US-10-117-722-290
Alignment Scores:
Pred. No.: 0 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservatives: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 17 Gaps: 2
US-09-671-687A-3 (1-949) x US-10-117-722-290 (1-2523)
QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
DB 45 GGAATAATAAGCTTTTATATTTTGTGTGGTGGACATGATTAACCTATTGGCAACTGG 104
QY 275 AspGlyArgPheAspGlyVal----LeuCysSerPheAlaCysValGluSerThrIleLeu 293
DB 105 GATGGAGAGATTGATGAGGTGACGCTTTGTAGTTTTCGCTGTGTTGAAAGTACAAATCTA 164
QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeu 313
DB 165 TTGCATCATCATGATATCATCCAGAGAGGTGACGCGAGAAAGGAGGCCCTCCCAACTT 224
QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333
DB 225 GCCTTTATGTCAAGAGGTGTGGGACAAAGTTTCATCCAGTCATATAAACCAGGCT 284
QY 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352
DB 285 ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAATTAATTTATACCTTAAT 344
QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysSerThrTrpTyrIleAspGlu 372
DB 345 GGCTCTTCTGTGACTCACAAACACCAATCCAAATCAAAATAATACATGGTACATTGATGAA 404
QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392
DB 405 GTTGCAGAGAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCA 464
QY 393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412
DB 465 CCACCACTCCAGCCTCCTCTGCTGTAACCTCACTACCCACCCAGAACAGATTCCTTTA 524
QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
DB 525 CCATTTCAGTCTCAACCAAGATGCCCAATACCAATGGAGATATTGGCCACAGTCCACTTCT 584
QY 433 LeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProPro 452
DB 585 CTGTACGCCCGAGTCTGTATGGAAGAGCTAAACACTGCACCCCGCTCCAGAGAGTCCACCC 644
QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
DB 645 TTGGCCATGCTCCTCTGGAACTCACATGTGTAGAGTGGGCTCATTTGGCTGAAGTTAAG 704
QY 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
DB 705 GAGAACCCCTCTTCTATGGGTAATCCGTGGATCGGTGATCGGTGAGCCAGGACTGAATGAA 764
```

```
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTGCTCGCTGGACTGGAACTGGAAAGATGAGTGTGCAGGCTGTACGATGGAACCTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
DB 825 GGCACCTCGGTATTTTACCTGTGCCCTGGAAGAGGCGCTGTTGTGAAACTGAAGAGTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGSCCTGACTCTAGGTTTGCATCATTTGCAGCGGTTTCCAAATCAGATTGAGCGCTGTAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTTGGAGGCTACTTAACTGAAGTAGTAGAAGAAATACTCCACCAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGAGGCTTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGGTTCATTACAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTACTTAGACTCAACTTATTCTGCTTATTCTGCTTTTAGTCTGTTCTGGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTTACTTAGACCCCAAGAAAGACGATGTAGAATATATATAGTGAACCAAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGACAGAAATTTGTAATCTCTGAGAAATATATGATATGTGTGTGCCACAAAATTT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAACACTGAGGAATAATCTTGAAGAGTGGAGGCTGCATCAGGATTTACTCTGAAGAA 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCCTGAGGAATTTCTTGAATATTTCTGTTTCATCATATTTTAAAGGTAGAACCTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAAAATTAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCGTTCCCACAATTCAGCAGTTGTAGAAATGGTCTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTGCAGAGGCAACCATCATGCTGATTTATTCAGATGCCTCGATT 1544
QY 753 GlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAAACTATTATAAAAAATTTTCTCTCTCTGGAATTTAAATATATACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTGAAGACACTCCACAGAGTCCGATATGTGGAGGCTTGCAGATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAGATGCTACCGATCCGACATCTCAGTCCGAAAAATCAAGCAGTTTGTGTAACACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATATAATATAACCCAGTGTCA 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGACTTTACCCGACTGGAGTGGAGACACGCGTGCATCCCTTGGCCAGAAATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
```

Db 1845 GAGTTATTGCTGTTCTCTGCATAGAAACAAGCCATGATGTTGCTTTTGTGAAGTATGGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsn 892
Db 1905 AAGGACGATTCTGCTGCTGCTCTTCTTGACAGCATGGCCGATCGGATGGTGTGAGAAAT 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGlnTyrlsMetSer 912
Db 1965 GGCCTTCAACATTCTCTCAAGTACCCCATGCCAGAAAGTAGGAGATCTTGAAGATGTCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgAtqIleGlnGlyCysAlaArgArgLeuLeu 932
Db 2025 CTGGAAGACCTGCAATCTCTTGACTCCAGGAGAAATCAAGGCTGTGCACGAAGACTGCTT 2084
QY 933 CysAspAlaTyrlsMetCysMetTyrlsGlnSerProThrMetSerLeuTyrls 949
Db 2085 TGTGATGCATATATGTGCATGTACACAGAGTCCAAACAATGAGTTGTACAAA 2135

RESULT 6

US-10-761-370-3
; Sequence 3, Application US/10761370
; Publication No. US20040219615A1
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: KOVALENKO, Andrei
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR R
; TITLE OF INVENTION: AND OTHER PROTEINS
; FILE REFERENCE: WALLACH=27
; CURRENT APPLICATION NUMBER: US/10/761,370
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US/09/646,403
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: IL 123758
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: PCT/IL99/00158
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: IL 126024
; PRIOR FILING DATE: 1998-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2116
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (691)..(691)
; OTHER INFORMATION: n is unknown.
US-10-761-370-3

Alignment Scores:
Pred. No.: 5,34e-261 Length: 2116
Score: 2468.00 Matches: 524
Percent Similarity: 91.21% Conservative: 5
Best Local Similarity: 90.34% Mismatches: 33
Query Match: 49.03% Indels: 22
DB: 18 Gaps: 6

US-09-671-687A-3 (1-949) x US-10-761-370-3 (1-2116)

QY 387 AspPheAspArgSerProLeuGlnProProValAsnSerLeuThrThrGlu 406
Db 15 GACTTTGACCGTTCTTACACACACATCCAGCCCTCTCTGTGAACACTACGACACCCGAG 74
QY 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426
Db 75 AACAGATTCCACTCTTTTACCATTCAGTCTCACCAAGATGCCAATACCAATGGAGATT 134
QY 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaPro 446
Db 135 GGCACAGTCCACTTCTCTCTGAGCCAGCTCTGTAATGAAGAGCTAAACACTGACACC 194

QY 447 ValGlnGluSerProLeuAlaMetProGlyAsnSerHisGlyLeuGluValGly 466
Db 195 GTCCAAAGAGAGTCCACCTTGGCCATGCCCTCTGGGAACCTACATGGTCTAGAAGTGGGC 254
QY 467 SerLeuAlaGluValLysGluAsnProProPheTyrlsValIleArgTrpIleGlyGln 486
Db 255 TCATTGGCTGAAGCTTAAGGAGAACCCCTCTCTATGGGGTAAATCCGTTGGATCGGTGAG 314
QY 487 ProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCys 506
Db 315 CCACCAAGGAGTGAATGAAGTCTCGCTGGAGCTGGAACCTGGAAGATGAGTGTGCAAGGCTGT 374
QY 507 ThrAspGlyThrPheArgGlyThrArgTyrlsPheThrCysAlaLeuLysLysAlaLeuPhe 526
Db 375 ACCGATGGAACTTCAGAGGCACTCGGTATTTACCTGTGCCCTGGAAGAGGCGCTGTTT 434
QY 527 ValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsn 546
Db 435 GTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTGCATCATTTGCAGCCGGTTTCCAAT 494
QY 547 Gln-IleGluArgCysAsnSerLeuAlaPheGlyGlyTyrlsSerGluValValGluGln 566
Db 495 CAAGATTGAGCGCTGAACCTCTTTAGCATTTGGAGGCTACTTAAAGTGAAGTAGT-GAAGA 553
QY 566 uAsnThrProProLysMetGluLysGluGluLeuMetIleGly---LysLysLys 585
Db 554 AAATACTCCACAAAAATGAAAAAAGAAAGCTTGGAGATAATGATGGGGAAGAAAGAAA 613
QY 585 sGlyIleGlnGlyHisTyrlsAsnSerCysTyrlsLeuAspSerThrLeuPhe-CysLeuPhe- 604
Db 614 GGCATCCAGGCGTCATTACAAATTTCTGKTACTTAGACTCAACCTTATTTCTGCTTATTTK 673
QY 605 AlaPheSerSerValLeu-AspThr---ValLeuLeuArgProLysGluLysAsnAsp-- 622
Db 674 GCTTTTAGTCTGCTCTINGACACTGGTGTACTTTAGACCCCAAGAAAAAGAAACGAT 733
QY 623 -ValGluTyrlsSerGluThrGlnGluLeuLeuArgThrGluLeuValAsnProLeuArg 642
Db 734 GTTAGAATATTTWKGMMACCCCAAGAGCTACTGAGGACAGAAATTTGTAATCTCTGAG 793
QY 642 gIleTyrlsTyrlsValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysVa 662
Db 794 AATATATGGATATGTGTGCCCAAAAAATTTAGAACTGAGGAAAAATTTGTAAGAGGT 853
QY 662 lGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIleLe 682
Db 854 GGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCTTGAGGAATTTCTTGAATATCT 913
QY 682 uPheHisHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysVa 702
Db 914 GTTTCATCATATTTTAAGGCTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGT 973
QY 702 lGlnAspCysTyrlsPheTyrlsGlnIlePheMetGluLysAsnGluLysValGlyValProTh 722
Db 974 ACAAGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAGAAAGTTGGCGTTCCAC 1033
QY 722 rIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaPr 742
Db 1034 AATTGACAGTTGTGTAAGATGGTCTTTTATCAACAGTAACCTGAAATTTGAGAGGCACC 1093
QY 742 oSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysIle 762
Db 1094 ATCATGCTCATTTATTCAGATGCTCGATTTGGAAGAGACTTTAAACTATTATTAATAATT 1153
QY 762 ePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysAr 782
Db 1154 -TTTCTCTCTGGAATTAGATATAACAGATTTACTTTGAAGACAC-CCGACAGCTGCGC 1211
QY 782 gIleCysGlyLeuAlaMetTyrlsGluCysArgGluCysTyrlsAspProAspIleSe 802
Db 1212 GATATGGAGGCGCTTCAATGTATGATGTAA-GAATGCTACGACGATCCGACACCCAG 1270
QY 802 rAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProly 822


```
Db 1271 C-----TGAAAAACAGCAGTTTGTAAACCTCAACACTCAAGTCACTTCATCCGAA 1327
QY 822 sArgLeuAsnHisLysTyAsnProValSerLeuProLysAspLeuProAsp---TrpAs 841
Db 1328 GAGGCTGAATCATATAATAACCCAGTGTACACTTCCCAAGACTTACCCGAGCTGGGAGA 1387
QY 841 pTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleG1 861
Db 1388 TTGGAGACAGCGCTGCATCCCTTGCAGAAATATGGAGTTATTTCGTCTCTGCATAGA 1447
QY 861 uThrSerHisTyValAlaPheValLysTyGlyLysAspAspSerAlaTrpLeuPhePh 881
Db 1448 AACAAACCACTATGTGCTTTTGTGAAGTATGGGAAGGAGATTCGCTCGCTCTCTCT 1507
QY 881 e-AspSerMetAlaAsp-ArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
Db 1508 TGGACAGCATGCCGATCCGGATGTGTGTCAGAAATGGCTCAACATTCCTCCCAAGTCMCC 1567
QY 901 ProCysProGluValGlyGluTyLeu-LysMetSer-LeuGluAsp---LeuHisSerL 919
Db 1568 CMTGCCCAAGACTAGGAGAGTACTTGGAAAGATGTCTCTGGAAGACCCCTGSAWTVCC 1627
QY 919 euAspSerArgArgIle-GlnGlyCysAlaArgArgLeuLeuCysAspAla-TyrMetCy 938
Db 1628 GGACTCCCAAGGAGAAATCCCAAGGCTGTGCACGAGACTGCTTTGTGATGCCATATATGTG 1687
QY 938 s-MetTy-Tyr-GlnSerProThrMetSerLeuTyLys 949
Db 1688 CCATGTACCAGAGTCCAAACATGAGTTTGTACAAA 1723
```

RESULT 7

```
US-10-264-237-204/c
; Sequence 204, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 204
; LENGTH: 1151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15)..(115)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1092)..(1092)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1113)..(1113)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1136)..(1136)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1142)..(1142)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-204
```

```
Alignment Scores: 2.65e-185 Length: 1151
Pred. No.: 1778.00 Matches: 347
Score: 90.89% Conservative: 2
Percent Similarity: 90.36% Mismatches: 15
Best Local Similarity: 35.32% Indels: 20
Query Match: 17 Gaps: 4
DB:

US-09-671-687A-3 (1-949) x US-10-264-237-204 (1-1151)

QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsn---GlnIleGluAtqCysAsn 552
Db 1150 CCTGACCCNAGGTNGTCATTTGCAGCCCGTTCANTCCAGATTGGAGCCGCTGTANC 1091
QY 553 SerLeuAla---PheGlyGlyTyLeuSerGluValValGluGluAsnThrPro---Pro 570
Db 1090 TCTTTAGCCATTGGGAGGGCTTACTTAAGTGAAGTAGTGGAGGAAATACTCCACCCAA 1031
QY 571 LysMetGluLysGluGlyLeuGluIleMetIle-GlyLysLysLysGlyIleGlnGlyHi 590
Db 1030 AAATGGGAAAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGAGGCATCCAGGTC 971
QY 590 sTyAsnSerCysTyLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerVal 610
Db 970 TTACAATCTTGTACTTAGACTCAACCTTATTCTGCTTATTTCCTTTAGTTCTGTCT 911
QY 610 uAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyTySerGluThrGl 630
Db 910 GGACACTGTGTACTTAGACCCAAAGAAAGACGATGTAGNATATATATAGTGAACCCA 851
QY 630 nGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyGlyTyValCysAlaTh 650
Db 850 AGAGCTACTTGAGGACAGAAATGTTAATCTCTGAGAAATATATGATATGTGTGTGCCAC 791
QY 650 rLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSe 670
Db 790 AAAAATTTATGAACTGAGGAAAAATCTTGAAGAGGTGGAGGCTGCATCAGGATTACCTC 731
QY 670 rGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGl 690
Db 730 TGAAGAAAAAGATCCTGAGGAATCTTGAATATATCTGTTTCATCATATATTTAAGGTAGA 671
QY 690 uProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyPheTyGlnIl 710
Db 670 ACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAT 611
QY 710 ePheMetGluLysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSe 730
Db 610 TTTTATGGAAAAAATGAGAAAGTTGGCGTTCCACAAATTCAGCAGTTGTTAGAAATGGTC 551
QY 730 rPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetPr 750
Db 550 TTTTATCAACAGTAACCTGAAATTTGCAGAGGACCAATCATGCTGATATTTCAAGATGCC 491
QY 750 oArgPheGlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIl 770
Db 490 TCGATTTGGAAAAAGACTTTAAACTATTTAAAAAATTTTCTCTCTCGGAATTAATAT 431
QY 770 eThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTy 790
Db 430 AACAGATTTTACTTTGAAGACACTCCACAGACAGTCCGATATGTGGAGGCTTGAAGTGA 371
QY 790 rGluCysArgGluCysTyTrpAspAspProAspIleSerAlaGlyLysIleLysGlnPheCy 810
Db 370 TGAGTGTAGAGAAATGCTTACGAYGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTG 311
QY 810 sLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyAsnPr 830
Db 310 TAAACCTGCAACACTCAAGTCCACCTTCATCCGAAGAGGCTGATATATAATAATAACCC 251
QY 830 oValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGl 850
Db 250 AGTGTCACTTCCCAAGAGACTTACCGAGCTGGAGACTGGAGACACAGCGCTGCATCCCTTGCCA 191
```


; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4253
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 517, 586
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-4253

Alignment Scores:
Pred. No.: 3.67e-111 Length: 617
Score: 1102.00 Matches: 203
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 2
Query Match: 21.89% Indels: 2
DB: 11 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-969-034-4253 (1-617)

QY	703	GlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThr	722
Db	616	CAAGATTGTTATTTCTATCAAAATTTTATGNA-AAAAATGAGAAAGTTGGCGTTCACCA	558
QY	723	IleGlnLeuLeuGluThrSer-PheIleAsnSerAsnLeuLysPheAlaGluAlaPr	742
Db	557	ATTCACAGTGTAGAAAGTCTTTTATCAACAGTAACNTGAAATTTGCAGAGGCACC	498
QY	742	oSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysI	762
Db	497	ATCATGTCTGATTATTTCAGATGCTCGATTGGAAAAGACTTTAAACTATTAAAAAAT	438
QY	762	ePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysAr	782
Db	437	TTTTCTCTCTCGAAATTAATAATAACAGATTTTACTTGAAGACACTCCACAGACAGTCCG	378
QY	782	gileCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSe	802
Db	377	GATATGTGGAGGCTTGCAATGTATGAGTGTAGAGATGCTACGACGATCCCGACATCTC	318
QY	802	xAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLy	822
Db	317	AGCTGGAAAATCAACAGTTTTTGTAAAACCTGCAACACTCAAGTCCACCTTCATCCGAA	258
QY	822	sArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTr	842
Db	257	GAGGCTGAATCATATAATAACCCAGTGTACCTTCCCAAAGACTTACCCGACTGGGACTG	198
QY	842	pArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluTh	862
Db	197	GAGACACGGCTGATCCCTTGCAGAAATATGGAGTTATTTGCTGTGTTCTCTGCATAGAAAC	138
QY	862	rSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPhePheSe	882
Db	137	AGCCCAATGTTGCTTTTGTGAAGTATGGGAAGGACGATTCGCTCGCTCTCTCTTGA	78
QY	882	pSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIleProGlnValThrProCy	902
Db	77	CAGCATGGCCGATCGGAGTGTGTGAGATGCGTTCAGAAATGCTTCAACATTCCTCAAGTCACCCCATG	18
QY	902	sProGluValGlyGlu	907
Db	17	CCGAGAGTAGGAGAG	2

RESULT 10
US-09-969-034-4254/c
; Sequence 4254, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4254
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 429, 468, 523
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-4254

Alignment Scores:
Pred. No.: 1.24e-86 Length: 557
Score: 879.00 Matches: 170
Percent Similarity: 91.44% Conservative: 1
Best Local Similarity: 90.91% Mismatches: 12
Query Match: 17.46% Indels: 5
DB: 11 Gaps: 2

US-09-671-687A-3 (1-949) x US-09-969-034-4254 (1-557)

QY	725	GlnLeuLeuGluTrpSerPheIleAsnSer---AsnLeuLysPheAlaGluAlaPro---	742
Db	557	CAGTGTAGATGGTTTTTATCAACAGTTAAACNTGAAATTTTGCAAAAGGCACCAT	498
QY	743	SeCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIle	762
Db	497	CAATGTTGATTATTCAGATGCTCGATT-NGAAAAGACTTTTAACTATTAAAAAAT	439
QY	763	PheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArg	782
Db	438	TTTCTCTTTTNGAAATTAATAATAACAGATTTTACTTGAAGACACTCCACAGACAGTCCGG	379
QY	783	IleCysGlyGlyLeu-AlaMetTyrGluCysArgGluCysTyrAsp-AspProAspIleS	802
Db	378	ATATGTGGAGGCTTTGCAATGTATGAGTGTAGAGATGTTACACCGATCCGACATCT	319
QY	802	erAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProL	822
Db	318	CAGCTGGAAAATCAAGCAGTTTTGTAAACCTTCAACACTCAAGTCCACCTTCATCCGA	259
QY	822	YsArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTr	842
Db	258	AGAGGCTGAATCATATAATAATAACCCAGTGTCTTCCCAAAGACTTACCCGACTGGGACT	199
QY	842	rpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluT	862
Db	198	GGAGACACGGCTGCATCCCTTGCAGAAATATGAGTTATTTGCTGTCTCTGCATAGAA	139
QY	862	hrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPhePheA	882
Db	138	CAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTTCTGCTGCTCTCTTTTG	79

```
QY      882  spSerMetAlaaspAaspGlyClyGlnAsnGlyPheAsnIleProGlnValThrProC 902
Db      78  ACAGCATGCCGATCGGATGGTGGTGCAGAAATGCTTCAACATTCCTCAAGTCACCCCAT 19
QY      902  ysProGluValGlyGlu 907
Db      18  GCCCAGAGTAGGAGAG 2

RESULT 11
US-09-983-965-1633
; Sequence 1633, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 1633
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 47-LIB2809-020-Q1-E1-D4
US-09-983-965-1633

Alignment Scores:
Pred. No.:      5,38e-59      Length:      376
Score:          626.00      Matches:    117
Percent Similarity: 94.35%      Conservative: 0
Best Local Similarity: 94.35%      Mismatches: 7
Query Match:      12.44%      Indels:     0
DB:               9          Gaps:      0

US-09-671-687A-3 (1-949) x US-09-983-965-1633 (1-376)

QY      703  GlnAspCysTyPheTyrglnIlePheMetGluLysAsnGluLysValGlyValProThr 722
Db      5   CAAGATAGTTACATCTATCAAAATTTTATGGAAGAAAATGAGAAAGTTGGAGTTCCTACA 64
QY      723  IleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaPro 742
Db      65  ATTCAGCAATTATTAGAAATGTTCTTTTATCAACAGTAACCTGAATTTGCTGAGGCACCA 124
QY      743  SerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIle 762
Db      125  TCATGCTCTGATTATTACAGATGCTCGATTTGGAAAAGACITTTAACTCTTTAAAAAAATT 184
QY      763  PheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArg 782
Db      185  TTTCTCTCTCTGGAACATAATAACAGATTATCTTGAAGACACCCCGGAGTGCCTCA 244
QY      783  IleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSer 802
Db      245  ATATGTGGAGGGCTTGCATATGATGAGTGATGAGATGCTATGATGACCTTGACATCTCA 304
QY      803  AlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisProLys 822
Db      305  GCTGGCAGCATCAAGCAGATTGTTGTAACCTGCAATGCTCAAGTCCACCTTCAGTCTAAG 364
QY      823  ArgLeuAsnHis 826
Db      365  AGGCTGAATCAT 376
```

```
RESULT 12
US-09-864-761-18145/c
; Sequence 18145, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18145
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007728.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EST HUMAN HIT: A1130924.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q10427, EVALUATE 1.70e-01
; OTHER INFORMATION: NT HIT: G111430965, EVALUATE 0.00e+00
US-09-864-761-18145
```

Alignment Scores:
Pred. No.: 3.38e-50 Length: 309
Score: 545.00 Matches: 102
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.83% Indels: 0
DB: 9 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-864-761-18145 (1-309)

QY	387	AspPheAspArgSerSerProProLeuGlnProProValAsnSerLeuThrGlu	406
DB	308	GACTTTGACCGTCTCTACCAACACCTCCAGCCTCTCTCTGTGAACCTACCTGACACCGAG	249
QY	407	AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle	426
DB	248	AACAGATTCCACTCTTTACCATTCAGTCTCACCAAGATGCCCAATACCAATGGAGATT	189
QY	427	GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuAsnThrAlaPro	446
DB	188	GGCCACAGCTCCACTTCTCTGTGAGCCAGTCTGTAATGGAAGAGCTAAACACTGTCACCC	129
QY	447	ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly	466
DB	128	GTCCAAGAGAGTCCACCTTTGCCATGCTCTCTCTGGGAACCTACATGGTCTAGAAAGTGGC	69
QY	467	SerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTTrpIleGlyGln	486
DB	68	TCATTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAATCCGTTGGATCGGTGAG	9
QY	487	ProPro 488	
DB	8	CCACCA 3	

RESULT 13

US-09-864-761-1386/c
; Sequence 1386, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35878
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(425)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-35878

Alignment Scores:
Pred. No.: 2,67e-35 Length: 425
Score: 412.00 Matches: 89
Percent Similarity: 63.12% Conservative: 0
Best Local Similarity: 63.12% Mismatches: 1
Query Match: 8.18% Indels: 52
DB: 10 Gaps: 1

US-09-671-687A-3 (1-949) x US-09-918-995-35878 (1-425)

QY 565 GluGluAsnThrProProLysMetGluLysGluGlyLeuGluMetIleGlyLysLys 584
Db 3 GAAGAAATACTCCACCAAAATGGAAGAGGCTTGAGATATATGATGGGAAGAG 62
QY 585 LysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPhe 604
Db 63 AAAGGCATCCAGGGTCATTACAAATCTTGTTACTTAGACTCAACCTTATTCGCTTATT 122
QY 605 AlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGlu 624
Db 123 GCTTTTAGTCTGTTCTTGACACTGTGTACTTAGACCCCAAGAAAGAACGATGTAGAA 182
QY 625 TyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIle--- 643
Db 183 TATTATAGTGAACCCAGAGCTACTGAGACAGAAATTTGTAATCTCTGAGAAAT-TAA 241
QY 643 ----- 643
Db 242 TTGATACTGGAACAAGACAGGTCAACCCGAAACTCTAAGGCATGTTAACACCTCCTTCC 301
QY 643 ----- 643
Db 302 TAGAAGAGAGGAAGATAAACATCTCTTGATCCTATTGTTGAGCCAGTTAAGAAATCAT 361
QY 644 -----TyrGlyTyrValCysAlaThrLysIleMet 653
Db 362 GTTAATCACACTTCTTACCCCTGTCCAGATATGGATATGTGTGCCCAAAAAATTATG 421
QY 654 Lys 654
Db 422 AAA 424

RESULT 15

US-10-152-319A-1764
; Sequence 1764, Application US/10152319A
; Publication No. US20040072160A1

GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Blashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1764
; LENGTH: 4847
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_021997
US-10-152-319A-1764

Alignment Scores:
Pred. No.: 7,68e-06 Length: 4847
Score: 160.50 Matches: 93
Percent Similarity: 32.76% Conservative: 40
Best Local Similarity: 22.91% Mismatches: 143
Query Match: 3.19% Indels: 130
DB: 17 Gaps: 16

US-09-671-687A-3 (1-949) x US-10-152-319A-1764 (1-4847)

QY 301 ProGluSerValThrGlnGluArg-----ArgProProLysLeuAlaPheMet 316
Db 272 CCTAGTGATCATCCATGCAGAACCCAGTGGCGCTGAAGCCCT----- 316
QY 317 SerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySer 336
Db 317 -----GGCGAGGGGGAGACACTCCAGCTCCGCTGGGTAGGCTCCATC----- 361
QY 337 ThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 362 -----GGGTCTGCCTCATCTTCTGTGTGCATCGGCGAGCGGCTCCAAAGAA 409
QY 357 AspSerGln-----ProGlnSerLysSerLysAsnThrTrpTyr 369
Db 410 GGGTCTCCCTGCACAAGCAGGCGTGGGCGCTCTCTCGGTGGAGCCCACT----- 463
QY 370 IleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe--- 388
Db 464 -----ACTGTCTCCGAGAAGCCAGGTCCAAAGCAGCCGAGGTGATGACTTCTTG 517
QY 389 -----AspArgSerSerProProLeuGlnProProValAsnSer 402
Db 518 GGAGACTTCGTGTGGCGAGAGGGTGTGGGTGAATGGCGTGAAGCCCGCGTGTGCAG 577
QY 403 LeuThrThrGluAsnArgPhe----- 409
Db 578 TACCTGGTGAGACACACAGTTTGGCCAGGCCAGTGGCGCGGTGTTGGATGATCCT 637
QY 409 ----- 409
Db 638 GTGGCAAGAACCATCGAGCTGTGGGTGGCTGCGCTACTTCGAGTGCACAGCTCTACAG 697
QY 410 -----HisSerLeuProPheSerLeuThrLysMetPro-----AsnThrAsnGlySer 425
Db 698 GGCATCTTACACGCGCTCCAAAGCTGACCCGCGCAGCCGCGAGCTGAGGGGCTCCGCGCAGT 757

```
QY 426 IleGlyHisSerProLeuSerAlaGlnSerValMetGluLeuAsnThrAla 445
Db 758 GACGGCCACTCAGTGAATCCCTCACCGCCCGAGACCTGTCCTCATTTCTGGCAGCGCC 817
QY 446 ProValGlnGluSerProLeu----- 453
Db 818 -----ACGCCCCCACTCACCGCCCGAGTATCCCCCTCGGGGAGAGTGTTC 865
QY 454 -----AlaMetProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluVal 471
Db 866 AACAGCTCCGTGAAGACGGCAATGAGTCTGGCTCCAACTCTCGGACAGTGGTCTGTG 925
QY 472 LysGluAsnProProPheTyr----- 478
Db 926 AAGCGTGTGACAGGACCTCCACCTGGGAGACCGTGTGCTGTGGTGGGACGAAGACC 985
QY 479 GlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGlu 498
Db 986 GGTGTGTGATGATGTGGGAGACGGACTTTGCCAAGGCGAGTGGTGTGGTGGAG 1045
QY 499 LeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThr 518
Db 1046 CTG---GACGAGCCCTTGGAGAGACGATGGGCGAGTGGCAGCGCACCGAGTACTTCCAG 1102
QY 519 CysAlaLeuLysLysAlaLeuPheValLysLeu----- 529
Db 1103 TGCCCAACCAAGTTCGTCTCTTTGACCAATCCACAAGTCCATCCGAATTGGCTTCCCA 1162
QY 530 -----LysSerCysArgProAspSerArgPheAla----- 539
Db 1163 TTTACAGTCCAGCCAGGCAAGAGACCAAGCGCATGGCTATGGGTGTCTCAGCCTTG 1222
QY 540 SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGly 557
Db 1223 ACCCAGAGCCCGAGAGTCTTCCATCAGCTCTGTGAGTCTGTGGCTCTCTCTGTGGT 1282
QY 558 GlyTyrLeuSerGlu-----ValValGluGluAsnThrProProLysMetGluLys 574
Db 1283 GGCGCGCGAGCGTAGTGGCTGTCTCACAGAGACCTTTTCAGCTATGCCCGGAAGATC 1342
QY 575 GluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyr-AsnSerCy 594
Db 1343 TCAGGCACCAACAGCCCTTGCAGGAGGACACTGAAGGAGAAACAGCAGCAGCATCGAACAGCTG 1402
QY 594 sTyrLeuAspSerThr 599
Db 1403 CTGGCTGAGCGTGA 1418
```

Search completed: April 18, 2005, 22:42:32
Job time : 1199 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:16:44 ; Search time 65 Seconds
(without alignments)
7476.356 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSPTWEERI.....RLLCDAYMCVQSPMTSLYK 949

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4971.5	98.8	956	1 CYLD HUMAN	Q9ncq7 homo sapien
2	4718.5	93.7	952	1 CYLD MOUSE	Q80tq2 mus musculu
3	4712	93.6	953	2 Q6H62	Q6h62 rattus norv
4	1207	24.0	778	2 Q70414	Q7q414 anopheles g
5	1136.5	22.6	551	2 Q8IPC3	Q8ipc3 drosophila
6	1136.5	22.6	639	2 Q8IPC5	Q8ipc5 drosophila
7	1134	22.5	550	2 Q8IPC4	Q8ipc4 drosophila
8	1132.5	22.5	639	2 Q8SYF0	Q8syf0 drosophila
9	1092.5	21.7	454	2 Q6TXJ6	Q6txj6 rattus norv
10	856	17.0	517	2 Q9VL04	Q9vl04 drosophila
11	619	12.3	1144	2 Q7JMS4	Q7jms4 caenorhabdi
12	183.5	3.6	258	2 Q7PUA3	Q7pja3 anopheles g
13	153.5	3.0	528	2 Q6BU87	Q6bu87 debaryomyce
14	151	3.0	653	2 Q6P5Z9	Q6p5z9 homo sapien
15	151	3.0	1427	1 REST HUMAN	P30622 homo sapien
16	149.5	3.0	1046	2 Q5S156	Q5s156 rattus norv
17	149	3.0	1046	2 Q9Z0H8	Q9z0h8 mus musculu
18	147.5	2.9	449	2 Q86WU4	Q86wu4 homo sapien
19	146	2.9	350	2 Q6PIA3	Q6pia3 homo sapien
20	145	2.9	547	2 Q96DZ5	Q96dz5 homo sapien
21	145	2.9	1556	2 Q9VRP1	Q9vrp1 drosophila
22	143.5	2.9	495	2 Q6FMR9	Q6fmr9 candida gla
23	143.5	2.9	1047	2 Q8EP81	Q8ep81 mus musculu
24	143	2.8	547	2 Q8MWL1	Q8mw11 homo sapien
25	142.5	2.8	1606	2 Q7PVQ8	Q7pvq8 anopheles g
26	142	2.8	1012	2 Q7T5I9	Q7tsi9 mus musculu
27	141	2.8	1012	2 Q8CHU1	Q8chu1 mus musculu
28	141	2.8	1429	2 Q6DFB6	Q6dfb6 xenopus lae
29	140.5	2.8	429	2 Q96C99	Q96c99 homo sapien
30	140	2.8	924	2 Q7VDY2	Q7vdy2 prochloroco
31	139.5	2.8	1921	2 Q86BR0	Q86br0 drosophila

ALIGNMENTS

RESULT 1

ID	CYLD_HUMAN	STANDARD;	PRT;	956 AA.
AC	Q9NCQ7; Q94934; Q7L3N6; Q96EH0; Q9NZX9;			
DT	25-OCT-2004 (Rel. 45, Created)			
DT	25-OCT-2004 (Rel. 45, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Probable ubiquitin carboxyl-terminal hydrolase CYLD (EC 3.1.2.15)			
DE	(Ubiquitin thiolesterase CYLD) (Ubiquitin-specific processing protease			
DE	CYLD) (Deubiquitinating enzyme CYLD) (HSPC057).			
GN	Name=CYLD; Synonyms=CYLD1, KIAA0849;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), DISEASE, AND TISSUE SPECIFICITY.			
RX	MEDLINE=20296617; PubMed=10835629; DOI=10.1038/76006;			
RA	Bignell G.R., Brown C., Biggs P.J., Lakhani S.R., Jones C., Hansen J.,			
RA	Blair E., Hofmann B., Siebert R., Turner G., Evans D.G.,			
RA	Schander-Stumpel C., Beemer F.A., Van Den Ouweland A., Halley D.,			
RA	Delpech B., Cleveland M.G., Leigh I., Leisti J., Rasmussen S.,			
RA	Wallace M.R., Fenske C., Banerjee P., Oiso N., Chaggar R., Merrett S.,			
RA	Leonard N., Huber M., Hohl D., Chapman P., Burn J., Swift S.,			
RA	Smith A., Ashworth A., Stratton M.R.;			
RT	"Identification of the familial cylindromatosis tumor suppressor			
RT	gene.";			
RL	Nat. Genet. 25:160-165(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Brain;			
RC	MEDLINE=99156230; PubMed=10048485;			
RX	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,			
RA	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. XII.			
RT	The complete sequences of 100 new cDNA clones from brain which code			
RT	for large proteins in vitro.";			
RL	DNA Res. 5:355-364(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Uterus;			
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,			
RA	Altschul S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Hopkins K.F., Mariani K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Diatchenko L., Ardalan K., Khouri A., Casavant T.L., Schetz T.E.,			
RA	Stapleton M.J., Udwin T.B., Toehiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			

Q757p1 ashbya goss
Q7rfu4 plasmodium
Q6a062 mus musculu
Q8BWD1 mus musculu
Q14527 homo sapien
Q922j3 mus musculu
Q6c1k8 varrowia li
Q9adn6 nicotiana t
Q9pvq6 anopheles g
Q6fqf0 candida gla
Q31152 neisseria m
Q7ddd8 neisseria m
Q9udt6 homo sapien
Q7z3n8 homo sapien

Db 121 KGLQIDVGPVKVQLSSEGEKPGVVRFRGPLLAERTVSGIFPGVLLBEGRQGGFTDGV 180
 QY 180 YGKQLFQCDCECG-FVALDKLELIEDDTALESDYAGPDTMQVLPPLPEINRSVSLKG 238
 Db 181 YGKQLFQCDCECGFVALDKLELIEDDTALESDYAGPDTMQVLPPLPEINRSVSLKV 240
 QY 239 GTTIESGTIVFCVLPKGSLSGYFVGVDNDPFGNWDGRFDGV-LCSFACVESTILLHN 297
 Db 241 GTTIESGTIVFCVLPKGSLSGYFVGVDNDPFGNWDGRFDGVQOLCSFACVESTILLHN 300
 QY 298 DIIP---ESVTOERRPPKPLAFMSRGVGDGSGSSHNKPKATGSDSPGNR-RSLFFYTLNG 353
 Db 301 DIIPALSESVTOERRPPKPLAFMSRGVGDGSGSSHNKPKATGSDSPGNRNSLFFYTLNG 360
 QY 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLOPPPVNSLTITNRRFHSLP 413
 Db 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLOPPPVNSLTITNRRFHSLP 420
 QY 414 FSLTKMPNTNGIGHSPLSLSAQSVMEELNTAPVQSPPLAMPNGSHGLEVGSIAEVKE 473
 Db 421 FSLTKMPNTNGIGHSPLSLSAQSVMEELNTAPVQSPPLAMPNGSHGLEVGSIAEVKE 480
 QY 474 NPPEFVGIRWIGOPPGCLNEVLAGLEDEACGCTDGTFRGTRYFTCALKKALFVKLSKR 533
 Db 481 NPPEFVGIRWIGOPPGCLNEVLAGLEDEACGCTDGTFRGTRYFTCALKKALFVKLSKR 540
 QY 534 PDSRFASLPQVSNQIRCNLSLAFGGVYLSEVVENTPPPKMEKEGLEIWKGGKGIQGHYNS 593
 Db 541 PDSRFASLPQVSNQIRCNLSLAFGGVYLSEVVENTPPPKMEKEGLEIWKGGKGIQGHYNS 600
 QY 594 CYLDSTLFCFLAFSSVLDTVLLRPKKNVYVYSETQELLRTIENPLRIYGVVCATKIM 653
 Db 601 CYLDSTLFCFLAFSSVLDTVLLRPKKNVYVYSETQELLRTIENPLRIYGVVCATKIM 660
 QY 654 KLRKILEKVEAASGFTSEKDEEFNLILFHHILRVEPLIKRSAGQVQDCYFYQIFME 713
 Db 661 KLRKILEKVEAASGFTSEKDEEFNLILFHHILRVEPLIKRSAGQVQDCYFYQIFME 720
 QY 714 KNEKVGVPITQQLLEWSFNSLNKFAEAPSLIIQMPRFQKPKFKIPPSLELNITDL 773
 Db 721 KNEKVGVPITQQLLEWSFNSLNKFAEAPSLIIQMPRFQKPKFKIPPSLELNITDL 780
 QY 774 LEDTPRQRCICGLAMCYECRECYDDPDISAGKIKQFCCTCNTQVHLHPKRLNHNKYNPVS 833
 Db 781 LEDTPRQRCICGLAMCYECRECYDDPDISAGKIKQFCCTCNTQVHLHPKRLNHNKYNPVS 840
 QY 834 PKDLFDWDRHGICPCQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNG 893
 Db 841 PKDLFDWDRHGICPCQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGGQNG 900
 QY 894 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCVMQSPMTSLYK 949
 Db 901 FNIPOVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCVMQSPMTSLYK 956

RESULT 2
 CYLD_MOUSE
 ID CYLD_MOUSE STANDARD; PRT: 952 AA.
 AC Q80Y02; Q80VB3; Q8BX23; Q8YL9; Q8CGB0;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Probable ubiquitin carboxyl-terminal hydrolase CYLD (EC 3.1.2.15)
 DE (Ubiquitin thiolesterase CYLD) (Ubiquitin-specific processing protease CYLD) (Deubiquitinating enzyme CYLD).
 GN Name=Cyld; Synonyms=Cyldi, Kiaa0849;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-620 FROM N.A.
 RP (ISOFORM 1).
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Hypothalamus;
 RX MEDLINE=22334683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato Y., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brucic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimson S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konasaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagahima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hahizume W., Imotani K., Ishii Y., Itoh M., Kigawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -I- FUNCTION: Negative regulator of TRAF2 and NF-kappa-B signaling
 CC pathway. Has deubiquitinating activity that is directed towards
 CC non-lys-48-linked polyubiquitin chains. The inhibition of NF-
 CC kappa-B activation is mediated at least in part, by the
 CC deubiquitination and inactivation of TRAF2 and, to a lesser
 CC extent, TRAF6 (By similarity).


```
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RG SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RG SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003628; AAN10740.1; -.
DR FlyBase; FBgn0032210; CG5603.
DR InterPro; IPR000938; CAP-Gly.
DR Pfam; PF01302; CAP_GLY; 1.
SQ SEQUENCE 550 AA; 62083 MW; C87C73651D88FD59 CRC64;

Query Match 22.5%; Score 1134; DB 2; Length 550;
Best Local Similarity 45.7%; Pred. No. 1.4e-65;
Matches 238; Conservative 87; Mismatches 168; Indels 28; Gaps 13;

QY 447 VOESPLAPPNGSHGLEVSLAEVKNP-----PFYGVIRWIGPPG-LNEVLAGLELED 501
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
38 IADYQPLIEIPGTE--LAIGSLVEV-SNFGVCEDLGVVRWIGIPGPGQKVLGVLEVED 94
QY 502 EC-----AGCTDGTFRGTRVFTCALKALFKVLKSCRPSRFPASLQP--VSNQIERCNLSA 555
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
95 ESNLKNVWASDGRHNGVRLFTCHDGRAIFVPANRCTADRRFADVDNSISANRVSSNHA 154
QY 556 FGYLSEVVEENTPPKM--EKGLEIMICKKGIQGHNSCYLDSLTCLFAPSSVLDTV 613
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155 FGADCPAIIYSGIPPLQIHNSDELASICGKFGIQQHNSCYLDATLFSMFTTFSVDSI 214
QY 614 LRPKEKNDVEYVSETOELLRTIEIYNPLRIYGVVCAIKMKLILEKVEAASGFTSEK 673
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
215 LYRRPGQDIRYSEVQKVRDEIYNPLKNVFRSDRVKMLRELLDQSSVSGLTCEEK 274
QY 674 DPEEFNLILFHHLIRVEPLKIRSAQKQVQDCYFYQFMKNEKVGVPITQQLLEWSFIN 733
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
275 DPEEFNLSLSQIMRVEPLKLSG---QDSFYQLFVEKEKLTLPVQQLFEQSFS 330
QY 734 SNLKFAEAPSLIIQMPREGKDFPKLPKTIIPSLLENTLLEDTTPQCRI CGSLAWYECR 793
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
331 SDIKLKEVPSCFIQMPREGKGYMYPRILPSQVLDVTDIENSPPQCSLCGLKLAIEYCR 390
QY 794 EGYDDPDISAG-KIQOFCKTCTQVHLHPKLNHNKYNPVSPLKDLPDMDWRHCCIFCQNM 852
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
391 DCFGSLQAGSGLECTAFCPKCLKTHSHIKRTHNVSKKIYSPKEFKIMA-EHMVVPRLYM 449
QY 853 ELFAVLCTIETSHYVAFVKYK--DDSAWLPFDSMADRDGGONGNIPOVTPCPBVGSLK- 910
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
450 ELFAVVCITETSHYVAFVKS GSGDPAPWCDFDSMADRKGEONGYNIPEITCVPELTQWSE 509
QY 911 ---MSLEDLHLSDRRIQGCARRLLCDAYMCMYQSTMSLY 948
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
510 ECARGINET-STNDKVLPEHAKRIFCDAYMCLYQSTDIMY 549

RESULT 8
Q8SYFO PRELIMINARY; PRT; 639 AA.
AC Q8SYF0;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE RE64280p.
GN ORFName=CG5603;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN SEQUENCE FROM N.A.
RP STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY071592; AAL49214.1; -.
DR FlyBase; FBgn0032210; CG5603.
DR InterPro; IPR000938; CAP-Gly.
DR Pfam; PF01302; CAP_GLY; 1.
SQ SEQUENCE 639 AA; 72162 MW; 6E29D3B8B757675B CRC64;

Query Match 22.5%; Score 1132.5; DB 2; Length 639;
Best Local Similarity 41.2%; Pred. No. 2.2e-65;
Matches 254; Conservative 103; Mismatches 203; Indels 57; Gaps 17;

QY 352 NGSSVDSQPSKSNWYIDEVAEDPAKSLTEI-STDFRSPPLQPPVNSLTENRPH 410
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 NHSAVDNQ-----HLEDV-----DLADILGTNWKRAAGPAAMLNNKSKTD---- 99
QY 411 SLPFSLTAKMNTNGSICHSPLSLSAQSVMEELNTPVQESPPPLAMPNGSHGLEVGLAE 470
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
100 -----PSNVSDILKPAFLKIEPEPELRFTIADYQPLIEIPGTE--LAIGSLVE 148
QY 471 VKENP---PFYGVIRWIGPPG-LNEVLAGLEDEEC---AGCTDGTFRGTRVFTCAL 521
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
149 V-SNPGVCEDLGVVRWIGIPGPGQKVLGVLEVEDSNLKNVWASDGRHNGVRLFTCHD 207
QY 522 KKALFKVLKSCRPSRFPASLQP--VSNQIERCNLSAGFGYLSVEVEENTPPKM--EKEL 577
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
208 GRAIFVPANRCTADRRFADVDNSISANRVSSNHAKEFGVADCPAIIYSGIPPLQIHNSDEL 267
QY 578 EIMIGKKGIQGHNSCYLDSLTCLFAPSSVLDTVLLRPKEKNDVEYVSETOELLRTIEI 637
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
268 ASICGKFGKIQQHNSCYLDSLTCLFAPSSVLDTVLLRPKEKNDVEYVSETOELLRTIEI 327
QY 638 VNPRLIYGVVCAIKMKLILEKVEAASGFTSEKDPESFLNLFHHLIRVEPLKIRS 697
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
328 VNPRLKNVFRSDRVKMLRELLDQSSVGLTCEEKPEEPNLSLSQIMRVEPLKLUSS 387
QY 698 AGQKQVQDCYFYQFMKNEKVGVPITQQLLEWSFINSLKFAEAPSLIIQMPREGKDFK 757
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
388 G---QDSFYQLFVEKEKLTLPVQQLFEQSFSHSDIKLKEVPSCFIQMPREGKGYK 443
QY 758 LFKKIPFSLLENTLLEDTTPQCRI CGGLAMVECRECYDDPDISAG-KIQOFCKTCTQV 816
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
444 MYPRILPSQVLDVTDIENSPPQCSLCGLKLAIEYECRDCFGSLQAGSGLECTAFCPKCLKT 503
QY 817 VHLHPKLNHNKYNPVSPLKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVKYK--DD 875
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
504 FHSIKRTHNVSKKIYSPKEFKIMA-EHMVVPRLYMELFAVVCITETSHYVAFVKS GSGPD 562
QY 876 SAWLPFDSMADRDGGONGNIPOVTPCPBVGSLK---MSLEDLHLSDRRIQGCARRLL 931
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
563 APWCDFDSMADRKGEONGYNIPEITCVPELTQWSEEGARSINET-STNDKVLPEHAKRI 621

932 LCDAYMCMYQSTMSLY 948
622 FCDAYMCLYQSTDIMY 638

RESULT 9
Q6TXJ6 PRELIMINARY; PRT; 454 AA.
AC Q6TXJ6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE LRRGT00003.
```


Db 371 KRSDDL-----LNLVCLDLFTKEEKLD---VWYKQCHDNRSRAKKSIRIKKIPP 418
QY 801 ISAGTKIQCKTCTNTQVHLHPKRLNHNKYNPVSUPKDLPL-----DWDWR 843
Db 419 VLSIQLKRF-----KHNHL-----NDTFSKIEIPIDIPFLNITKYTSASMSPELGD-- 466
QY 844 HGCIPQNMELFAVLG-----IETSHVAVPVKYKDDSAWLPFD 882
Db 467 -----KIFELFALVCHIGSVNTGHVIVIKNG--NGQWLKFD 501

RESULT 14
Q6P5Z9 PRELIMINARY; PRT; 653 AA.
AC Q6P5Z9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE RSN protein (Fragment).
GN Name=RSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.243603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klauener R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC062543; AAH62543.1; -
DR InterPro; IPR000938; CAP_Gly.
DR Pfam; PF01302; CAP_Gly; 2.
DR PROSITE; PS00845; CAP_Gly_1; 2.
DR PROSITE; PS0245; CAP_Gly_2; 2.
FT NON_TER 653
SQ SEQUENCE 653 AA; 4074A22377D4BB17 CRC64;

Query Match 3.0%; Score 151; DB 2; Length 653;
Best Local Similarity 20.8%; Pred. No. 0.39;
Matches 99; Conservative 55; Mismatches 171; Indels 157; Gaps 20;

QY 260 GYFVGVDMPNIGWDRFGVLCSFACVBSSTILLHNDIIPESVQERRPPKLAFWSRG 319
Db 85 GQWAGIVDEPIGKNDGVSAGVR-YQCEP-----LKGIF-----TRPSKL---TRK 127
QY 320 VGDGKSSSHNPKATGSDPDGNRRSEFLFYTLNGSSVDSPQSKSNWTWIDEVAEDPAK 379
Db 128 V-----QAEDEANGLQTTASRATSPLCSTSTASWSSSPSPSN---IPQKPSQPA- 175

QY 380 SLTEISTDFDRSSPPLOPPPVNSLTNTENPHSLPSPSILTKMPTNGSICHSPLSLSAQSV 439
Db 176 -----AKEPSATPISNLT-----KTASESIS 197
QY 440 EELNTAPVQSPPLAMPNGSHGLVEGSLAEVKENPPFYGVIRWIGOPGLNEVLAGLESL 499
Db 198 NLSEAGSIKK-----CERELKIGDRVLVGTTKA--GVVRLGETDFAKGEWCGLV 246
QY 500 EDECAGCTDGTFRGTFTYTCALKKALFKVLKSCRPDSRFASLQPVSNQIERNSLAFGGY 559
Db 247 -DEPLGKNDGAVAGTRYFQCPKYGLFA-----PVHKVTIKGF--- 283
QY 560 LSEVVEENTPPKMEKGL-FIMIGKKGIQGHVNSCYLDSTLFLCLFAESSVDTVLLRPK 618
Db 284 -----PSTTPAKAKANAVRMATTSASLKRSPSASSLS-----MSSVASSVSSRP- 330
QY 619 EKNDVEYSETQELLRTIENVLRIVGYVCATKMK-----LRKILEKVEAAGSF 668
Db 331 -----SRTGLTETSSRYARKISGTTALQELAKKQKHIEQLLAERDLERAFAKA- 381
QY 669 TSEKQPEEFNLIL-----FHHIL-----RVEPLLKIRSAGQKVQDC 705
Db 382 TSHVGEIEQELALARDGHDQHVLEAKMDQLRTMVEADREKVELLNQLNSEEKRVEDL 441
QY 706 YF 707
Db 442 QF 443

RESULT 15
REST HUMAN STANDARD; PRT; 1427 AA.
ID REST HUMAN
AC P30622;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 03-JUL-2004 (Rel. 44, Last annotation update)
DE Restin (Cytoplasmic linker protein-170 alpha-2) (CLIP-170) (Reed-
DE Sternberg intermediate filament associated protein).
GN Name=RSN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripherial blood monocytes;
RX MEDLINE=92289675; PubMed=1600942;
RA Bilbe G., Delabie J., Brueggen J., Richener H., Asselbergs F.A.M.,
RA Cerletti N., Sorg C., Odink K., Tarcsay L., Wiesendanger W.,
RA de Wolf-Peters C., Shipman R.;
RT "Restin: a novel intermediate filament-associated protein highly
RT expressed in the Reed-Sternberg cells of Hodgkin's disease.";
RL EMBO J. 11:2103-2113(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92405160; PubMed=1356075; DOI=10.1016/0092-8674(92)90240-D;
RA Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
RT "CLIP-170 links endocytic vesicles to microtubules.";
RL Cell 70:887-900(1992).
CC -I- FUNCTION: Seems to be a intermediate filament associated protein
CC that links endocytic vesicles to microtubules.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic; associated with the
CC cytoskeleton.
CC -I- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P30622-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P30622-2; Sequence=VSP_000765;
CC -I- TISSUE SPECIFICITY: Highly expressed in the Reed-Sternberg cells
CC of Hodgkin's disease.
CC -I- SIMILARITY: Contains 2 CAP-Gly domains.

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:25:35 ; Search time 1059 Seconds
(without alignments)
5304.848 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSGGLWSQKVTSPWEERI.....RLLCDAYMCVQSPTMSLYK 949

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp
-Q/cgn2_1/USPTO_spool_p/US09671687/runat_18042005_115017_17115/app_query.fasta_1.1095
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CGN_1_1_708 @runat_18042005_115017_17115 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: N_Geneseq_16Dec04:*
- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4983	99.0	3302	12	Adq95917 T cell ac
2	4983	99.0	6831	10	Adc24898 Human bre
3	4980	99.9	3302	12	Adq95915 T cell ac
4	4971.5	98.8	5371	10	Abv75394 Human CYL
5	4971.5	98.8	5371	13	Adrl4488 Human NF-

6	4968.5	98.7	3311	12	Adq95919	Adq95919 T cell ac
7	4942	98.2	4527	3	AAa08589	AAa08589 Human cyt
8	4601	91.4	2845	3	AAh18625	AAh18625 Human cDN
9	3876.5	77.0	4716	4	AAa41617	AAa41617 CDNA enco
10	3676	73.0	2523	4	AAi58410	AAi58410 Human pol
11	3676	73.0	2523	5	Ado98620	Ado98620 DNA enco
12	3676	73.0	2523	9	Adb48380	Adb48380 Novel hum
13	3523	70.0	4286	4	AAi60196	AAi60196 Human pol
14	3329	66.1	2488	4	AAa41081	AAa41081 CDNA enco
15	3022.5	60.0	2569	4	AAh18478	AAh18478 Human cDN
16	2468	49.0	2116	2	AAz07515	AAz07515 Human RAP
17	1778	35.3	1151	6	ABl89642	ABl89642 Human pol
18	1494	29.7	1013	6	ABl89643	ABl89643 Human pol
c	1490.5	29.6	1160	5	AAa90524	AAa90524 DNA enco
c	1262	25.1	837	4	AAH07749	AAH07749 Human cDN
20	1215	24.1	758	4	AAH05084	AAH05084 Human cDN
21	1102	21.9	617	6	ABQ60558	ABQ60558 Human col
22	936	18.6	5433	4	ABl09704	ABl09704 Drosophil
23	879	17.5	557	6	ABQ60559	ABQ60559 Human col
c	867.5	17.2	8990	4	AAK77783	AAK77783 Human imm
25	856	17.0	1655	4	ABl09705	ABl09705 Drosophil
27	626	12.4	376	8	ABX51704	ABX51704 Bovine ES
28	568	11.3	318	12	ADK71948	ADK71948 Human ori
29	545	10.8	303	12	ADK71946	ADK71946 Human ori
c	545	10.8	309	4	AAI20692	AAI20692 Probe #10
c	545	10.8	309	4	ABa65741	ABa65741 Human foe
c	545	10.8	309	4	AAI45907	AAI45907 Probe #14
c	545	10.8	309	4	ABa47848	ABa47848 Human bre
c	545	10.8	309	4	ABa32825	ABa32825 Probe #11
c	545	10.8	309	4	AAK39883	AAK39883 Human bon
c	545	10.8	309	4	AAK14143	AAK14143 Human bra
c	545	10.8	309	5	AAI06390	AAI06390 Probe #63
c	545	10.8	309	6	ABs13981	ABs13981 Human gen
c	514	10.2	288	12	ADK71950	ADK71950 Human ori
c	512	10.2	483	4	AAi11479	AAi11479 Probe #14
c	512	10.2	483	4	ABa53149	ABa53149 Human foe
c	512	10.2	483	4	AAI32754	AAI32754 Probe #14
c	512	10.2	483	4	ABa42721	ABa42721 Human bre
c	512	10.2	483	4	ABa22920	ABa22920 Probe #13
c	512	10.2	483	4	AAK36850	AAK36850 Human bon

ALIGNMENTS

RESULT 1
ADQ95917
ID ADQ95917 standard; cDNA; 3302 BP.
XX
AC ADQ95917;
DT 07-OCT-2004 (first entry)
XX
DE T cell activation associated cDNA #48.
XX
KW ss; gene; antiallergic; antiarthritic; antidiabetic;
KW anti-HIV; antimicrobial; antiarthritic; immunosuppressive;
KW neuroprotective; gene therapy; T cell activation; diagnosis;
KW autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;
KW diabetes; allergic disease; infectious disease; AIDS; chronic rejection;
KW organ; bone-marrow transplant.
XX
OS Homo sapiens.

XX	Key	Location/Qualifiers
XX	CDS	243..3104
XX	FT	/*tag= a
XX	XX	
PN	XX	W02004058805-A2.
XX	XX	15-JUL-2004. - <i>late</i>
XX	XX	25-DEC-2003; 2003WO-JP016715.


```

PR 26-DEC-2002; 2002JP-00376365.
PR 27-DEC-2002; 2002US-0436473P.
PR 25-APR-2003; 2003JP-00122113.
PR 28-APR-2003; 2003US-0465792P.
PR 21-OCT-2003; 2003JP-00360559.
PR 22-OCT-2003; 2003US-0512846P.
XX
PA (ASAH-) ASahi KASEI PHARMA CORP.
XX
PI Matsuda A, Yoneta S;
XX
XX WPI: 2004-593134/57.
DR P-PSDB; ADQ95918.
XX
PT New purified protein involved in T cell activation, useful for
PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
PT and infectious diseases.
XX
PS Claim 4; SEQ ID NO 95; 2828pp; English.
XX
CC The invention relates to purified proteins and genes encoding them, that
CC are involved in T cell activation (I) and has an amino acid deletion,
CC substitution or addition in the amino acid sequences. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
CC marrow transplant. This sequence corresponds to a cDNA for a protein
CC involved in T cell activation.
XX
SQ Sequence 3302 BP; 980 A; 648 C; 773 G; 901 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 3302
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 12 Gaps: 4
US-09-671-687A-3 (1-949) x ADQ95917 (1-3302)
QY 1 MetSerSerGlyLeuTrpSerGlnGluValThrSerProTyrTrpGluGluArgIle 20
DB 243 ATGAGTTTCAGGCTTATGGACCCAGAAAGTCACTTCACCTTCTGGGAGAGCGGATT 302
QY 21 PheTyrIleuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 303 TTTTACTTGCTTCTCAAGAAATGACGGTTACAGACAAACACAAAGCTCCTTAA 362
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisArgIle 60
DB 363 GTACCGAAGGAAAGTAGTAGACAGTATATTCAGATCGTTCTGTGGGCAATCAAGGATT 422
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLysIleLeuGlnGlnProHisAla 80
DB 423 CCTTCTGCAAAAGCAAGAAATTCAGATTGGATTAAATTTCTAGAGCAACTCATGCA 482
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
DB 483 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 542
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerIleuPheLysAsnArgAsnArgLeuSer 119
DB 543 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCGCTGTTTAAAAACAGAAAACAGACTAAGT 602
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 603 AAAGGCTCCAAATAGACGTGGGCTGCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 662
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB

```

```

663 AAATTTCTCGGAGTTGTACGCTTCAGAGGACCCCTGTTACGAGAGGACAGTCTCCGGA 722
160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
723 ATATTTCTTTGGAGTTGAAATGCTGGAAGAAGGTCGTGGTCAAGGTTTCTACTGACGGGGTG 782
180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
783 TACCAAGGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTGATTGGAC 842
199 LysLeuGluLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
843 AGCTAGAACTCATAGAAGATGATGACACTGCACTTGGAAAGTGAATACGAGGTCCTGGG 902
219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
903 GACACATGCAAGTCCGAACTTCTCTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT 962
239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
963 CGAGAAACAATAGAACTCTGGAAACAGTTATATCTGTGATGTTTTCAGGAAAAAAGC 1022
259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
1023 TTAGGATATTTTGTGGTGTGGACATGGATAACCTTATTGGCAACTGGGATGGAAGATT 1082
279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
1083 GATGGAGTCAGCTTTGTAGTTTGGTGTGTTGAAAGTACAATTTCTATTGCACTCAAT 1142
298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
1143 GATATCATCCAGAGAGTGTGCGCAGGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1202
318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
1203 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 1262
338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
1263 TCAGACCTTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGCTCTTGTT 1322
357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
1323 GACTCACACCAACCAATCCAAATCAAAATAATACATGTTGATGATGAAGTTGCAGAAAGAC 1382
377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
1383 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCTCCAG 1442
397 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
1443 CCTCTCTCTGTGAACCTACACGACGAGAACAGATTCCACTCTTTACCATTCAGTCTC 1502
417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
1503 ACCAAGATGCCAATACCAATGGAGATTGTTGGCCACAGTCCACTTCTCTGTACGCCAG 1562
437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
1563 TCTGTAAATGGAAGAGCTTAAACACTGCACCGCTCCAGAGAGTCCACCGCTTGGCCATGCT 1622
457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
1623 CCTGGGAACCTCACATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAAACCCCTCT 1682
477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
1683 TTCTATGGGTATATCCGTTGGATCGGTACGCCACCAAGAGTGAATGAAGTCTCGCTGGA 1742
497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
1743 CTGGAACTGGAAGATGATGTGAGGCTGTACGAGTGAACCTTTCAGAGGACCTCGGTAT 1802

```

QY 517 PheThrCysAlaIleuLysIysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1803 TTCACCTGTCCTGAAGAGCGCTGTTGTGAAACTGAAGAGCTGCAGCGCTGACTCT 1862
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1863 AGTTTGCATCATTTGACGCGGTTTCATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1922
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 1923 CGAGGCTACTTAAGTCAAGTAGTAGAAGAAATFACCCACAAAATGGAAAAAGAGGC 1982
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1983 TTGGAGATTAATGATGGGAGAGAGAAAGGCATCCAGGGTCATTACAAATCTTGTGTACTTA 2042
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2043 GACTCAACCTTATTTCTGCTTATTTGCTTTTGTAGTTCTGTCGACACTGTGTACTTAGA 2102
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2103 CCCAAAGAAAGAACGATGAGATATATAGTGAACCCCAAGAGCTACTGAGACAGAA 2162
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 2163 ATTGTTAATCCTCGAGAATATATGATATGTGTGCCACAAAAATTTATGAACCTGAGG 2222
QY 657 LysIleLeuGluLysValGluAlaLysSerGlyPheThrSerGluGluLysAspProGlu 676
DB 2223 AAAATACTTTGAAAAGGTGAGGCTGCATCAGGATTTTACCTCTCAAGAAAAAGATCTCGAG 2282
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
DB 2283 GAATCTTGAATATCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAAAATAAGA 2342
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2343 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGGAATAAATGAG 2402
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
DB 2403 AAAGTTGGCGTTCCCAATTCACAGATTGTTAGAAATGGTCTTTTATCAACAGTAACCTG 2462
QY 737 LysPheAlaGluAlaProSerCysLeuIleLeGlnMetProArgPheGlyLysAspPhe 756
DB 2463 AAATTTGCAGAGCCACCATCATGCTGATTATTCAGATGCTCGATTGTGAAAAGACTTT 2522
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2523 AAATATTAAAAAATTTTCTCTCTCGAATTAATAATATACAGATTACTTTGAAGAC 2582
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2583 ACTCCAGACAGTCCGCGATATGCGAGGCTTGCAATGTATCAGTGTAGAGAATGCTAC 2642
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2643 GACGATCCGGACATCTCAGCTGGAAAAATCAACAGTTTTTGTAAAAACCTGCAACACTCAA 2702
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2703 GTCCACTTCATCCGAAGAGGCTGATCATATAATATAACCATGTGCTACTTCCCAAGAC 2762
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 2763 TTACCGGACTGGGACTGGACACGCGTGCATCCCTTGCAGAAATATGAGATTATTGCT 2822
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
DB 2823 GTTCTCTGATAGAAAACGCACTATGTTGCTTTTGTGAAGTATGGGAGGACGATTCT 2882

QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 2883 GCTGCTCTCTTTTTCACGATGCGGATGGTGGTCAATATGGCTTCAACAT 2942
QY 897 ProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeu 916
DB 2943 CCTCAAGTCAACCCCATGCCAGAAAGTAGGAGTACTTTGAAGATGTCTCTGGAAGACCTG 3002
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
DB 3003 CATTCCTTGGACTCCAGGAAATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3062
QY 937 MetCysMetTyrClnSerProThrMetSerLeuTyrLys 949
DB 3063 ATGTGCATGTACCAGAGTCCAAATGAGTTTGTACAAA 3101

RESULT 2
ADC24898
ID ADC24898 standard; cDNA; 6831 BP.
XX
AC ADC24898;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human breast specific nucleic acid (BSNA) DEX0238_49, SEQ ID NO:49.
XX
KW Human; breast specific polypeptide; BSP; breast specific nucleic acid;
KW BSNA; breast cancer; metastasis; non-cancerous disease; breast tissue;
KW identification; monitoring; diagnosis;
KW engineered breast tissue production; transgenic animal; drug screening;
KW cytostatic; gene therapy; vaccine; chromosome 16p13.3; gene; ss.
XX
OS Homo sapiens.
XX
PN WC02003020900-A2.
XX
PD 13-MAR-2003.
XX
PF 29-AUG-2002; 2002WO-US027777. - *late*
XX
PR 31-AUG-2001; 2001US-0316306P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Liu C, Salceda S;
XX
DR WPI; 2003-290182/28.
DR P-PSDB; ADC24816.
XX
PT New breast specific polypeptide useful for identifying, diagnosing,
PT monitoring, staging, imaging and treating breast cancer and non-cancerous
PT disease states in breast.
XX
PS Example 1; SEQ ID NO 49; 264pp; English.
XX
CC The invention relates to breast specific polypeptides (BSPs) and nucleic
CC acids (breast specific nucleic acids; BSNA) encoding them. The invention
CC also relates to vectors and host cells comprising a BSNA sequence;
CC antibodies against BSNA; the recombinant production of BSNA; methods of
CC detection of BSNA or BSNA in a sample; kits for detecting a risk of
CC cancer or presence of cancer in a patient; and vaccines comprising a BSNA
CC or BSP. The invention additionally discloses fragments, mutants, fusion
CC proteins, homologous proteins and allelic variants of BSNA; methods for
CC identifying and designing agonists and antagonists of BSNA; methods for
CC identifying and monitoring breast tissue; producing engineered breast
CC tissue for treatment and research; producing transgenic animals and cells
CC comprising BSNA sequences; aptamers evolved to bind specifically to BSNA;
CC and single exon probes based on BSNA sequences. BSNA, BSNA and
CC antibodies against BSNA are useful for identifying, diagnosing,
CC monitoring, staging, imaging and treating breast cancer (including breast
CC cancer metastases), and non-cancerous disease states in breast tissue.
CC BSNA and BSNA may additionally be used to identify and monitor breast
CC tissue, in screening for BSP agonists and antagonists, and in the

CC production of engineered breast tissue for treatment or research. BSNAs
 CC may also be used in gene therapy and in the production of transgenic
 CC animals and cells. The present sequence represents a breast specific
 CC nucleic acid (BSNA) used in an example of the invention.

XX
 SQ Sequence 6831 BP; 1987 A; 1256 C; 1494 G; 2094 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 6831
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 10 Gaps: 4

US-09-671-687A-3 (1-949) x ADC24898 (1-6831)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
 DB 1585 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCCCTTACTGGGAAGCGGATT 1644
 QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 1645 TTITACTTCTCTTCAAGAATCGAGGTTACAGACAAACAAACAAAGCTCCTTAAA 1704
 QY 41 ValProLysGlySerIleGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 DB 1705 GTACCGAAGGAGTATAGACAGTATATTCAGATCGTTCCTGGGGCATTCAGGATT 1764
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 DB 1765 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGAATTAATAATTTAGAGCAACCTCATGCA 1824
 QY 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeuLeu 99
 DB 1825 GTTCTCTTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 1884
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 1885 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCAGCTGTTTAAACACAGAACACACTAAGT 1944
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 1945 AAAGGCTTCAAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 2004
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 2005 AAATTTCTCGAGTTGTACGCTTTCAGAGGACCCCTGTTCAGCAGAGGACAGTCTCCGGA 2064
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 DB 2065 ATATTTCTTTGGAGTTGAAATTTGCTGGAAAGAGGTCGTGGTCAAGTTTCACTGACGGGGTG 2124
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 2125 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTTCATTGGAC 2184
 QY 199 LysLeuLeuLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 2185 AAGCTAGAACTCTAGAAGATGATGACACTGCACTTGGAAAGTGAATACGCAAGTCTCTGGG 2244
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 2245 GACACAAATGACGTGACGACTTCTCTTTCGAAATAAATCCAGAGTTTCTTTGAAGTT 2304
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 2305 GGAGAAACAATAGAACTCGAACAGATTATATCTGTGATGTTTGTCCAGGAAAGAAAGC 2364
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 2365 TTAGGATATTTTGTGTGGAGCATGGATAACCCCTATTGGCACTGGGATGGAAGATT 2424

QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 2425 GATGGAGTCAGCTTTGTAGTTTGGTGTGTGTTGAAAGTACATTTCTATTGGCATCAAT 2484
 QY 298 AspilelleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
 DB 2485 GATATCATCCAGAGAGTGTGACGAGGAAGAGGCGCTCCCAAACTTGGCTTTATGTCA 2544
 QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 2545 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 2604
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 DB 2605 TCAGACCTCTGGAATAGAAAAACAGATCTGAATTTATATACCTTAAATGGGTCTCTGT 2664
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
 DB 2665 GACTCAACACCAATCCAAATCAAAAAATACATGTTGATGAAGTTGCAGAAAGAC 2724
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
 DB 2725 CTGCAAAATCTTACAGAGATATCTACAGACTTTGACCGTTCTTCAACCACTCCAG 2784
 QY 397 ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 DB 2785 CCTCTCTCTGTAACCTCACTGACCAACGAGAACAGATTCACCTCTTTACCATTCAGTCTC 2844
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLysLeuSerAlaGln 436
 DB 2845 ACCAAGATCCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTCTCTGTGAGCCAG 2904
 QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 2905 TCTGTAATGGAGAGCTAAACACTGCACCCCTCAAGAGAGTCCACCTTGGCCATCGCT 2964
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 2965 CCTGGGAACCTCATGCTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCT 3024
 QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 3025 TTCTATGGGGTAAATCCGTTGGATCGGTGACCCACAGGACTGAATGAAGTGTCTCGTGA 3084
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 DB 3085 CTGGAACTGGAAGATGAGTGTGCGCTGTACGGATGGAACCTTCAGAGGACCTCGGTAT 3144
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
 DB 3145 TTCACCTGTGCCCTGAAGAAGGCGCTGTTGTGAACTGAAGAGCTGCAGGCTGACTCT 3204
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 DB 3205 AGGTTTGCATCATTTGACGCGGTTTCCAATTCAGATTGAGCGCTGTAACTCTTTAGCATTT 3264
 QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
 DB 3265 GGAGGCTACTTAAGTAGAGTAGAAGAAAAATCTCCACAAANAATGGAAAAAGAGGC 3324
 QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
 DB 3325 TTGGAGATATGATTGGGAAGAAAGAGCATCCAGGCTCATTAACAATTTCTTTACTTTA 3384
 QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
 DB 3385 GACTCAACCTTATTTCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTGA 3444
 QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGlnLeuLeuArgThrGlu 636
 DB 3445 CCCAAGAAAGAACGATGATGATATATATGTAAGAACCCCAAGAGCTACTGAGGACAGAA 3504
 QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656

Db 3505 ATGTTAATCTCTGAGATATATGATATGCTGTGTCACAAAATATGAACCTGAGG 3564
 Qy 657 LysIleLeuGluLysValGluAlaIleSerGlyPheThrSerGluGluLysAspProGlu 676
 Db 3565 AAAATACTTTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTCTGAAGAAAAGATCTCTGAG 3624
 Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
 Db 3625 GAATCTTGAATATCTTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAATAAAGA 3684
 Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
 Db 3685 TCAGCAGGTCAAAAGGTACAGATTTGTACTTCTATCAAAATTTTATGGAAAAATGAG 3744
 Qy 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTyrPheSerPheIleAsnSerAsnLeu 736
 Db 3745 AAAGTTGGGGTTCCTCCCAATTCAGCAGTTGTTAGAAATGGTCTTTTATCAACAGTAACCTG 3804
 Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
 Db 3805 AAATTTGCAGAGGCACCATCATCTCGATTATTCAGATGCTCGATTGGAAAAGACTTT 3864
 Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
 Db 3865 AAATTAATTTAAATAATTTTCTCTCTGGAATTAATAATAACAGATTTACTTTGAAGAC 3924
 Qy 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
 Db 3925 ACTCCAGACAGTCGCGGATATGTGAGGGCTTGCAATGTATGAGTGTAGAGAATGCTAC 3984
 Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
 Db 3985 GACGATCCGACATCTCAGCTGGAAAATCAAGCAGTTTTTGTAAAACCTGCAACACTCAA 4044
 Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
 Db 4045 GTCCACCTTCATCCGAAGAGGCTGGAATCATAAATAAATCAACAGTGCACCTTCCCAAAGAC 4104
 Qy 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
 Db 4105 TTACCCGACTGGAGCTGGAGACACGCTGCATCCCTTGCAGAAATATGGATTTATTTGCT 4164
 Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 Db 4165 GTTCTCTGCATAGAAAACGACCTATGTGCTTTTGTGAAGTATGGGAAGGACGATCT 4224
 Qy 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
 Db 4225 GCCTGGCTCTCTTTGACAGCATGGCGCATCGGGATGGTGGTCAGAAATGGCTTCAACATT 4284
 Qy 897 ProGlnValThrProCysProGluValGlyLysLeuLysMetSerLeuGluAspLeu 916
 Db 4285 CCTCAAGTCAACCCATGCCAGAAAGTAGAGAGTACTTTGAAGATGCTCTGGAAAGACCTG 4344
 Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 Db 4345 CATTCCTTGGACTCCAGGAGATCCCAAGGCTGTGCACGAGAGACTGCTTTGTGTGATAT 4404
 Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrIlys 949
 Db 4405 ATGTGCATGTACCAAGATCCAAACATCAGTTGTACAAA 4443
 RESULT 3
 ID ADQ95915 standard; cDNA; 3302 BP.
 XX AC ADQ95915;
 XX AC ADQ95915;
 DT 07-OCT-2004 (first entry)
 XX T cell activation associated cDNA #47.
 DE
 XX

ss; gene; anti-allergic; antiarthritic; antiasthmatic; antidiabetic;
 KW anti-HIV; antimicrobial; antirheumatic; immunosuppressive;
 KW neuroprotective; gene therapy; T cell activation; diagnosis;
 KW autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;
 KW diabetes; allergic disease; infectious disease; AIDS; chronic rejection;
 KW organ; bone-marrow transplant.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 CDS 243..3104
 FT /*tag= a
 XX
 PN W02004058805-A2.
 XX
 PD 15-JUL-2004.
 XX
 PP 25-DEC-2003; 2003WO-JP016715.
 XX
 PR 26-DEC-2002; 2002JP-00376365.
 PR 27-DEC-2002; 2002US-0436473P.
 PR 25-APR-2003; 2003JP-00122113.
 PR 28-APR-2003; 2003US-0465792P.
 PR 21-OCT-2003; 2003JP-00360559.
 PR 22-OCT-2003; 2003US-0512846P.
 XX
 PA (ASAH-) ASahi Kasei PHARMA CORP.
 XX
 PI Matsuda A, Yoneta S;
 XX
 DR WPI; 2004-593134/57.
 DR P-PSDB; ADQ95916.
 XX
 PT New purified protein involved in T cell activation, useful for
 PT diagnosing, preventing and/or treating acquired immunodeficiency
 PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
 PT and infectious diseases.
 XX
 PS Claim 4; SEQ ID NO 93; 2828bp; English.
 XX
 CC The invention relates to purified proteins and genes encoding them, that
 CC are involved in T cell activation (1) and has an amino acid deletion,
 CC substitution or addition in the amino acid sequences. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
 CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
 CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
 CC -marrow transplant. This sequence corresponds to a cDNA for a protein
 CC involved in T cell activation.
 XX
 SQ Sequence 3302 BP; 981 A; 648 C; 772 G; 901 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 3302
 Score: 4980.00 Matches: 947
 Percent Similarity: 99.48% Conservative: 1
 Best Local Similarity: 99.37% Mismatches: 1
 Query Match: 98.93% Indels: 4
 DB: 12 Gaps: 4
 US-09-671-687A-3 (1-949) x ADQ95915 (1-3302)
 Qy 1 MetSerSerGlyLeuTyrPheSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
 Db 243 ATGATTTTCAGCTTATGGAGCCAGAAAAGTCACTTCACTTCTGGAAAGACGGATT 302
 Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 Db 303 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACACAAACAAACAAAGCTCTTAAA 362
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 Db 363 GTACCCGAAGGAAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCGCATTCAGGATT 422

late

QY	61	ProSerAlaIleGlyLysLysGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80	QY	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436
Db	423	CCTTCTGCAAAAGGCAAGAAATCAGATTGATTAATAAATCTAGAGCAACCTCATGCA	482	Db	1503	ACCAAGATCCCAATACCAATGGAGTATTTGGCCACAGTCCACTTCTTCTGTGAGCCAG	1562
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99	QY	437	SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro	456
Db	483	GTCTCTTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT	542	Db	1563	TCGTGTAATGGAGAGCTAAACACTGACCCGCTCAAGAGAGTCCACCTTGGCCATGCCT	1622
QY	100	LeuAlaIleThrAsnGlyGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119	QY	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476
Db	543	TTGGCAATTTACCAATTTGTGAGAGAGTTTCAGCTGTTTTTAAACAACAGACACTAAGT	602	Db	1623	CCTGGGAACCTCAGATGCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT	1682
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139	QY	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly	496
Db	603	AAAGGCTCCAAATAGACGTGGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA	662	Db	1683	TTCTATGGGGTATTCGGTTGGATCGGTACGCCACAGGACTGAATGAAGTCTCGCTGGA	1742
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159	QY	497	LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516
Db	663	AAATTTCTTGGAGTTGAAATGCTGGAGAGAGTCTGGTCAAGGTTTCACTGACGGGGTG	782	Db	1743	CTGGAACCTGAAGATGAGTGTGAGGCTGTACGATGGAAACCTTCAGAGGCACCTCGGTAT	1802
QY	160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal	179	QY	517	PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer	536
Db	723	ATATTTCTTGGAGTTGAAATGCTGGAGAGAGTCTGGTCAAGGTTTCACTGACGGGGTG	782	Db	1803	TTACCTGTGCTTGAAGAGGCGCTGTTGTGAACTGAAGAGCTGCAGGCTGACTCT	1862
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198	QY	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556
Db	783	TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC	842	Db	1863	AGGTTTGCATCATTCGACGCGGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCATTT	1922
QY	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218	QY	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProLysMetGluLysGluGly	576
Db	843	NAAGTAGAACTCATAGAGATGATGACACTGCATTCGAAAGTGATTACGCGAGTCTCTGG	902	Db	1923	GGAGGCTACTTAAGTAGTAGAAGAAATACTCCCAAAATAATGGAAAAAGAGGC	1982
QY	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238	QY	577	LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596
Db	903	GACACAATGCAGTCAAGCTTCTCTCTTGGAAATAAATCCAGAGGTTTCTTTGAAGGTT	962	Db	1983	TTGAGATATATGATTGGGAAGAAAGAGGATCCAGGCTCATTACAAATCTTGTACTTA	2042
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258	QY	597	AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg	616
Db	963	GGGAACAACATAGAACTGGAACAGATTATATCTGTGATGTTTCCAGAGAAAGAAC	1022	Db	2043	GACTCAACCTTATCTGCTTATTTGCTTTTGTGTTTCTGTTCTGGACACTGTGTTACTAGA	2102
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278	QY	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
Db	1023	TTAGGATATTTTGTGTGGACATGATGATAACCTTATTTGGCAACTGGGATGGAAGATT	1082	Db	2103	CCCAAGAAAAAGAACCATGATAGTAATATATAGTGAAACCCCAAGAGCTACTGAGACAGAA	2162
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297	QY	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656
Db	1083	GATGGAGTGCAGCTTTGTAGTTTGTGCTGTGTTGAAAGTACAAATTTCTATTGCAATCAAT	1142	Db	2163	ATTGTTAATCTCTGAGATATATGATATATGATATGTTGTGCCACAAAAATATGAACTGAAG	2222
QY	298	AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer	317	QY	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
Db	1143	GATATCATCCAGAGAGTGTGACGAGAAAGAGGCTCCCAAACTTGGCTTTATGTCA	1202	Db	2223	AAAATCTTCAAAAGGTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAAGATCCTGAG	2282
QY	318	ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr	337	QY	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696
Db	1203	AGAGGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACCAGGCTACAGGATCTACC	1262	Db	2283	GAATTTCTGTAATATCTGTTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGA	2342
QY	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerVal	356	QY	697	SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu	716
Db	1263	TCAGACCTGGAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGCTCTCTGTT	1322	Db	2343	TCAGCAGGTCAAAAGGTGACAGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG	2402
QY	357	AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp	376	QY	717	LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu	736
Db	1323	GACTCAACCAACCAATCCAAATCAAAAAATACATGTTGATGATGAGTTGACAGAGAC	1382	Db	2403	AAAGTTGGCGTTCCCAATTCAGCAGTTGTTAGATGGTCTTTTATCAACAGATAACCTG	2462
QY	377	ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln	396	QY	737	LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe	756
Db	1383	CTTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACTCCAG	1442	Db	2463	AAATTTTCAGAGGCACCATCATGTCTGATTATTTCAGATGCTCGATTTTGGAAAAAGACTTT	2522
QY	397	ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu	416	QY	757	LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp	776
Db	1443	CCTCCTCTGTGAACCTCACTGACCACCGAGAACAGATTCCACTCTTTTACCATTTCAGTCTC	1502	Db	2523	AAACTATTTAAAAAAATTTTTCCTTCTCTGGAATTTAAATATAACAGATTTTACTTGAAGAC	2582
				QY	777	ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr	796

Db 2593 ACTCCAGACAGTGGCGGTATGTGGAGGGCTTGCATGTATGAGTGTAGAGATCTTAC 2642
 Qy 797 AspAspProAspIleSerLaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
 Db 2643 GACGATCCGACATCTCAGCTGGMAAAATCAAGCAGTGTGTAAACCTGCAACACTCAA 2702
 Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuPheLysAsp 836
 Db 2703 GTCCACCTTCATCCGAAGAGGCTGAATCAATAATATAACCCAGTGTCTCCCAAGAC 2762
 Qy 837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
 Db 2763 TTACCCGACTGGAGCTGGAGACAGCGCTGCATCCCTTGCAGANAATGAGTATTGTGT 2822
 Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 Db 2823 GTTCTCTGCATAGAACCAAGCCACTATGTGTCTTTGTGAAGTATGGGAAGCAGATTCT 2882
 Qy 877 AlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
 Db 2883 GCCTGGCTCTTCTTTGACAGCATGGCCGATCGGGATGGTGGTCAAGATGGCTTCAACATT 2942
 Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 Db 2943 CCTCAAGTCACCCCATGCCAGAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 3002
 Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 Db 3003 CATTCCTTGGACTCCAGGAGATCCAGAGCTGTGCACGAGACTGCTTTGTGATGATAT 3062
 Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 3063 ATGTGCATGTACAGATCCAAATGAGTGTGTACAAA 3101

RESULT 4

ABV75394
 ID ABV75394 standard; DNA; 5371 BP.
 XX AC ABV75394;
 XX DT 18-MAR-2003 (first entry)
 XX DE Human CYLD polypeptide encoding DNA.
 XX KW CD40; NF-kappaB essential modulator; NEMO; CYLD; immunosuppressive;
 XX KW antiinflammatory; cytostatic; gene therapy; human; gene; ds.
 XX OS Homo sapiens.

Key Location/Qualifiers
 CDS 392..3262
 FT /*tag= a
 FT /product= "CYLD"

WO200292761-A2.

21-NOV-2002.

07-MAY-2002; 2002WO-US014570.

08-MAY-2001; 2001US-00851673.

(IMMUNEX) IMMUNEX CORP.

Derry JMJ, Fanslow WC, Dougall WC;

WPI; 2003-120669/11.

P-PSDB; ABB82783.

Identifying compounds that alter one or more biological activities of
 CD40 by modulating the binding of NEMO and CYLD, useful for treating
 disorders of the immune system, and inflammatory and cancer diseases.

XX Example 5; Page 38-44; 48pp; English.
 XX The invention relates to methods of identifying compounds that alter one/
 CC more biological activities of CD40. One method involves screening for
 CC molecules that modulate the binding of NF-kappaB essential modulator
 CC (NEMO) and CYLD. The methods and compositions of the invention of
 CC determining compounds that agonize or antagonize a CD40 signaling
 CC activity, are useful for the further definition of CD40-mediated
 CC signaling pathways, and for manipulation of CD40-mediated cellular
 CC responses. They also provide therapeutic agents for treating disorders of
 CC the immune system, and inflammatory and cancer diseases. The present
 CC sequence represents a DNA encoding the human CYLD polypeptide
 XX Sequence 5371 BP; 1600 A; 950 C; 1183 G; 1638 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 5371
 Score: 4971.50 Matches: 948
 Percent Similarity: 99.16% Conservative: 0
 Best Local Similarity: 99.16% Mismatches: 1
 Query Match: 98.76% Indels: 7
 DB: 10 Gaps: 5

US-09-671-687A-3 (1-949) x ABV75394 (1-5371)

Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
 Db 392 ATCAGTTTCAGGCTTATGGAGCCAAAGAAAAGTCACTTCACTTCTGGGAAGAGCGGATT 451
 Qy 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 Db 452 TTTTACTTGTCTTCAAGAAATGACGGTTTACAGACAAACAAACACAAAGCTCTTAAA 511
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGluAspArgSerValGlyHisSerArgIle 60
 Db 512 GTACCGAAGGAAAGTAGTAGACAGTATATTCAGATCGTCTGTGGGCGCATTCAGGATT 571
 Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 Db 572 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAATCTAGAGCAACCTCATGCA 631
 Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 Db 632 GTTCTCTTGTGATGAAAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACTT 691
 Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 Db 692 TTGGCAATTACCAATTGTGTAGGAGAGGTTTCAGGCTGTTTAAAAACAGAAAACAGACTAAGT 751
 Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 Db 752 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGNAAGTACAGCTGAGATCTGGGGAGAA 811
 Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 Db 812 AAATTTCTCGAGTGTGTAGCTTTCAGAGGACCCCTGTGTAGCAGAGAGACAGTCTCCGA 871
 Qy 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 Db 872 ATATTCTTTGGAGTTGAATTGCTGGAAGAAGGTTCGTGCTCAAGGTTTTCACGTGACGGGGTG 931
 Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 Db 932 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGGTGTTTGTTCATTTGAC 991
 Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 Db 992 AAGCTAGAACTCATAGAGATGATGACACTGTCATTGGAAAGTATTACGACAGTCTGG 1051
 Qy 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 Db 1052 GACACAATGCGAGTCCGAACCTTCTCTCTTGGAAATAAACTCCAGAGTTTCTTTGAAGGTT 1111

QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258	QY	594	CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal	613
Db	1112	CGAGAAACAATAGAAATCTGAAACAGTTATTTCTGTGATGTTTTCGCCAGAAAGAAAGC	1171	Db	2192	TGTACTTAGACTCAACCTTATTTCTGCTTATTTGCTTTTAGTCTGTTCTGTCACACTGTG	2251
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278	QY	614	LeuLeuArgProLysGlyLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu	633
Db	1172	TTAGGATATTTTGTGTGTGACATGGATACCCCTATTGGCAACTGGGATGGAGATTT	1231	Db	2252	TTACTTTAGACCCCAAGAAAGAACGATGTAGAAATATTATTAGTGAAACCCCAAGAGCTACTG	2311
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297	QY	634	ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet	653
Db	1232	GATGAGTGCAGCTTTGTAGTTTGGTGTGTGAAAGTACAAATCTATTGGACATCAAT	1291	Db	2312	AGGACAGAAATGTTTAATCCTCTCGAGATATATGATATATGTTGTGCCACAAAAATATG	2371
QY	298	AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla	314	QY	654	LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys	673
Db	1292	GATATCATCCAGCTTTATCAGAGAGTGTGCGCAGAAAGGAGGCTCCCAAACTTGCC	1351	Db	2372	AAACTGAGGAAATATCTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAA	2431
QY	315	PheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThr	334	QY	674	AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu	693
Db	1352	TTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACAAGGCTACA	1411	Db	2432	GATCCTGAGGAATCTTTGAATATCTGTTTCATCATATTTTAAGGGTAGAACCTTTTGCTA	2491
QY	335	GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly	353	QY	694	LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu	713
Db	1412	GGATCTACCTCAGACCTCGAAATAGAAACAGATCTGAATTTATTATACCTTAAATGGG	1471	Db	2492	AAATTAAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATCGAA	2551
QY	354	SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTptTyrIleAspGluVal	373	QY	714	LysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsn	733
Db	1472	TCTCTGTTGACTCAACACCAATCCAAATCAAAATAATACATGTTGATGAAGTT	1531	Db	2552	AAAAATGAGAAAGTTGGCGTCCCACAATTCAGCAGTTGTAGAAATGGCTTTTATCAAC	2611
QY	374	AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerPro	393	QY	734	SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly	753
Db	1532	GCAGAAGACCTCGCAAAATCTTTACAGAGATATCTACAGACTTTGACCGTCTTCACCA	1591	Db	2612	AGTAACCTCAAAATTTTCAGAGGACCATCATGTCTGATTATCAGATGCCCTCGATTGGA	2671
QY	394	ProLeuGlnProProValAsnSerLeuThrGluAsnArgPheHisSerLeuPro	413	QY	754	LysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeu	773
Db	1592	CCACTCCAGCCTCCTCTGTGAACCTCACTGACCACCGAGAACAGATTCCACTCTTACCA	1651	Db	2672	AAAGACTTTAAACTATTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTA	2731
QY	414	PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu	433	QY	774	LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg	793
Db	1652	TTTCAGTCTCACCAAGATGCCAATACCAATGGAAATTTGGCCACAGTCCACTTTCTCTG	1711	Db	2732	CTTTGAAGACACTCCACAGACAGTCCCGATATGTGGAGGGCTTGCAATGTATGAGTAGA	2791
QY	434	SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu	453	QY	794	GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys	813
Db	1712	TCAGCCAGCTCTGTAATGGAAGAGCTAAACACTGCACCCGCTCAGAGAGTCCACCCCTG	1771	Db	2792	GAAATGCTACGACCATCCGACATCTCAGCTGGAAAAAATCAAGCAGATTTTGTAAACCTGC	2851
QY	454	AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu	473	QY	814	AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu	833
Db	1772	GCCATGCCCTCTGGGAACTCACATGCTCTAGAAAGTGGGCTCATTTGCTGAAGTTAAGGAG	1831	Db	2852	AAACCTCAAGTCCACCTTCATCCGAAAGAGGCTGAATCATTAATAATAACCCAGTGTCACTT	2911
QY	474	AsnProProPheTyrGlyValIleArgTptIleGlyGlnProProGlyLeuAsnGluVal	493	QY	834	ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu	853
Db	1832	AACCTCTCTTTATGGGTAAATCCGTGGATCCGTCAGCCACAGCATGAATGAAGTG	1891	Db	2912	CCCAAGAGCTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTCCGCAATATGGAG	2971
QY	494	LeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGly	513	QY	854	LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys	873
Db	1892	CTCGCTGGACTGGAATGGAAGATGAGTGTGCAAGGCTGTACGGATGGAACCTTCAGAGGC	1951	Db	2972	TTATTTGCTGTTCTCTGTCATAGAAACAAAGCCACTATGTTGCTTTTGTGAAGTAGTGGGAAG	3031
QY	514	ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArg	533	QY	874	AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly	893
Db	1952	ACTCGGTATTTCACCTGTGCCCTGGAAGAGCGCTGTTTGTGAACACTGAAGAGCTGCAGG	2011	Db	3032	GAGATTCGCCCTGGGCTCTTCTTTGACAGCATGCCCATCGGGATGGTGGTGCAGAAATGGC	3091
QY	534	ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer	553	QY	894	PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu	913
Db	2012	CCTGACTCTAGGTTGGATCAITGGACCGGTTTCCAATCAGATTGAGCGCTGTAACTCT	2071	Db	3092	TTCAACATCTCTCAAGTCAACCCATGCCAGAGAGTAGAGAGTACTTTGAAGATGTCTCTG	3151
QY	554	LeuAlaPheGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu	573	QY	914	GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCys	933
Db	2072	TTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAGAAATAACTCCACCAAAAAATGGAA	2131	Db	3152	GAAGACCTGCTATCTCTGGACTCCAGGAGAAATCCAGGCTGTGCACGAGACTGCTTTGT	3211
QY	574	LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer	593	QY	934	AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys	949
Db	2132	AAAGAAGGCTTGGAGATATGATTGGGAAGAAAGGCGCATCCAGGGGTCTATTCAATCT	2191	Db	3212	GATGCATATATGTCATGTATCCAGAGTCCAAACAAATGAGTTTGTACAAA	3259

RESULT 5

ADRI14488

ID ADRI14488 standard; DNA; 5371 BP.

XX ADRI14488;

AC ADRI14488;

XX 21-OCT-2004 (first entry)

XX Human NF-kappaB pathway-associated gene SeqID489.

DE

XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;

XX antiarthritic; antirheumatic; gastrointestinal-gen; antiaesthetic;

XX antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;

XX immunosuppressive; vulnary; gene therapy; immune disorder; immune

XX inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;

XX hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;

XX hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;

XX X-linked anhidrotic ectodermal dysplasia; immunodeficiency;

XX viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;

XX viral replication; host cell survival; evasion of immune response;

XX rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;

XX atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;

XX autoimmune disorder; hyper immune activity;

XX aberrant acute phase response; hypercongenital condition; birth defect;

XX necrotic lesion; wound; organ transplant rejection;

XX aberrant signal transduction; proliferating disorder; cancer;

XX HIV propagation; gene; ds; human.

XX Homo sapiens.

OS

XX WO2004065577-A2.

PN

XX 05-AUG-2004.

XX 13-JAN-2004; 2004WO-US000798.

XX 14-JAN-2003; 2003US-0440068P.

PR 12-MAY-2003; 2003US-0469757P.

PR

XX (BRM) BRISTOL-MYERS SQUIBB CO.

XX Nadler SG, Neubauer MG, Feder JN, Carman J;

PI WPI; 2004-562168/54.

XX P-PSDB; ADRI14489.

DR

XX New isolated polynucleotides and polypeptides associated with NF-kappaB

XX pathway, useful for diagnosing, treating, or preventing disorders or

XX diseases associated with NF-kappaB pathway.

XX Claim 1; SEQ ID NO 489; 237pp; English.

PS

XX This invention relates to the novel association of protein sequences (and

XX the genes which encode them) to the NF-kappaB pathway. The invention may

XX be useful for the production of compounds with an antiinflammatory,

XX cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,

XX gastrointestinal-gen, antiaesthetic, antiatherosclerotic,

XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or

XX vulnary activity or for gene therapy. The proteins and nucleotides are

XX useful for diagnosing, preventing, treating, or ameliorating conditions

XX or diseases associated with the NF-kappaB pathway. The condition is an

XX immune disorder, an inflammatory disorder, an inflammatory disorder

XX related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,

XX hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM

XX syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic

XX ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,

XX hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell

XX survival, evasion of immune responses, rheumatoid arthritis, inflammatory

XX bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick

XX syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper

XX immune activity, disorders related to aberrant acute phase responses,

XX hypercongenital conditions, birth defects, necrotic lesions, wounds,

XX organ transplant rejection, conditions related to organ transplant

XX rejection, disorders related to aberrant signal transduction,

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human gene which is
 CC subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.

XX
 XX Sequence 5371 BP; 1600 A; 950 C; 1183 G; 1638 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 5371
 Score: 4971.50 Matches: 948
 Percent Similarity: 99.16% Conservative: 0
 Best Local Similarity: 99.16% Mismatches: 1
 Query Match: 98.76% Indels: 7
 DB: 13 Gaps: 5

US-09-671-687A-3 (1-949) x ADRI14488 (1-5371)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
 DB 392 ATGAGTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCACCCCTACGGAGAGCGGATT 451
 QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 452 TTTTACTTGCTTCTTCAGAAATGCGAGCTTACAGCAACAAACACAAAAGCTCCTTAAA 511
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
 DB 512 GTACCGAAGGGAAGTAGACAGATATATTCAGATCGTCTTCTGTGGGCATTCAGGATT 571
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 DB 572 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATTCCTAGAGCAACCTCATGCA 631
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 DB 632 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGATTACAGAGATTACTT 691
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 692 TTGGCAATTACCAATTGTGGAGAGAGGTTCCAGCTGCTGTTTAAAAACAGAAACAGACTAAGT 751
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuAArgSerGlyGluGlu 139
 DB 752 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 811
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 812 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 871
 QY 160 IlePhePheGlyValGluLeuLeuGluGluValArgGlyGlnGlyPheThrAspGlyVal 179
 DB 872 ATATTTCTTTGGAGTTGAATTCCTGGAAAGAGGTCGTGCTCAAGGTTTCATGACGGGGTG 931
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 932 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGGTGTTTGTTCATTTGGAC 991
 QY 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 992 AAGCTAGAACTCATAGAAGATGATGACACTTGGCAATTTGAAAGTGTATTACGCAAGGCTCTGGG 1051
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 1052 GACACAAATGAGTTCGAACTTCCTCTCTTGGAAATAAACTCCAGAGTTCTTTTGAAGGTT 1111
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 1112 GGAGAAACAATAGAACTGGAAACAGATTATATTCTGTGATGTTTTCAGGAGAAAGAAAGC 1171
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProLysGlyAsnTrpAspGlyArgPhe 278
 DB 1172 TTAGGATATTTTGTGTGGAGCATGATAACCTTATTTGGCAACTGGGATGGAAGATT 1231


```
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1232 GATGGAGTCAGCTTTGTAGTTGTGGTGTGTTGAAAGTAGTACATTTCTATTTCACATCAAT 1291
QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
DB 1292 GATATCATCCAGCTTTATCAGAGAGTGTGACGAGAAAGGAGGCTCCCAAACTTGCC 1351
QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
DB 1352 TTTATGTCAAGAGGTGTGGGACAAAGGTTCATCCAGTCATATTAACCAAGGCTACA 1411
QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheThrLeuAsnGly 353
DB 1412 GGATCTACCTCAGACCTCGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGG 1471
QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyIleAspGluVal 373
DB 1472 TCTTCTGTGTGACTCACACCAACCAATCCAAATCAAAAAATACATGGTACATTTGATGAAGTT 1531
QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
DB 1532 GCAGAAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTCACCA 1591
QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
DB 1592 CCACTCCAGCTCCTCTGTGAACCTCACTGACCCAGAGAACAGATTCACACTCTTTTACCA 1651
QY 414 PheSerLeuThrLysMetProLanThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
DB 1652 TTCAGTCTCACCAGAGTGCCTCAATACCAATGGAAGTATTTGGCCACAGTCCACTTCTCTG 1711
QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu 453
DB 1712 TCAGCCAGTCTGTAATGAAGAGCTAAACACTGCACCCGTCGACAGAGTCCACCCCTTG 1771
QY 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
DB 1772 GCCATGCTCTCTGGGAACCTCACATGCTGTAGAAGTGGCTCATTTGGCTGAAGTTAAGGAG 1831
QY 474 AsnProProPheThrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
DB 1832 AACCCCTCCTTTCTATGGGGTAAATCCGTTGGATCGGTGAGCCAGCCAGGACTGAAATGAAGTG 1891
QY 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
DB 1892 CTCGCTGGGACTGGAATGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGAGGC 1951
QY 514 ThrArgTyThrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
DB 1952 ACTCGGTATTTACCTGTGCTGGAAGAGCGCTGTTGTGAAACTGAAGAGCTGCAGG 2011
QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
DB 2012 CCTGACTCTAGGTTTGATCATTTGACGCGGTTTCCAATCAGATTGAGCGCTGTAACCTCT 2071
QY 554 LeuAlaPheGlyGlyTyThrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
DB 2072 TTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATAACTCCACCAAAAATGGAA 2131
QY 574 LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyAsnSer 593
DB 2132 AAAGAAGGCTTGGAGATATATGATTGGGAAGAAAGGCAATCCAGGGTCATTACAATCT 2191
QY 594 CysTyThrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
DB 2192 TGTTACTTAGACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTG 2251
QY 614 LeuLeuArgProLysGluLysAsnAspValGluTyThrSerGluThrGlnGluLeuLeu 633
DB 2252 TTACTTTAGACCCCAAGAAAGACGATGTAGATAATATTATAGTGAAACCCCAAGAGCTACTG 2311
```

```
QY 634 ArgThrGluIleValAsnProLeuArgIleTyGlyTyValCysAlaThrLysIleMet 653
DB 2312 AGACAGAAATTTGTTAATCTCTGAGATATATGGAATATGTTGTGTGTCACAAAAATATG 2371
QY 654 LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
DB 2372 AAATCTGAGGAATACTTGAAGAGTGGAGGCTGCATCAGATTTTACCTCTGAAGAAAAA 2431
QY 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
DB 2432 GATCTCTGAGGAATTTCTTGAATATTTCTTTCATCATATTTTAAAGGTAGAACCTTTGCTA 2491
QY 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyThrPheTyGlnIlePheMetGlu 713
DB 2492 AAAATTAAGATCAGCAGGTCAAAAGGTACAAGATTTGTACTTCTATCAAAATTTTATGGAA 2551
QY 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn 733
DB 2552 AAAAATGAGAAAGTTGGCGTTCCCAATTCACAGATTGTAGAAATGGTCTTTTATCAAC 2611
QY 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
DB 2612 AGTAACCTGAATTTGCAGAGGACCATCATGTCTGATTAATTCAGATGCGCTCGATTGGA 2671
QY 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
DB 2672 AAAGACTTTAAACTATTTAAAAAAATTTTCTCTCTGGAATTTAAATATAACAGATTTA 2731
QY 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyThrGluCysArg 793
DB 2732 CTTGAAGACACTCCCAAGACAGTCCCGATATGTGGAGGGCTTGCAATGTATGATGTAGA 2791
QY 794 GluCysTyThrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
DB 2792 GAATGCTACAGCAATCCGACATCTCAGCTGGAAAAATCAAGCAGATTTGTTGTAACCTGC 2851
QY 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyThrAsnProValSerLeu 833
DB 2852 AACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATAACCCAGTGTCACTT 2911
QY 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
DB 2912 CCAAAAGACTTACCCGACTGGGACTGGAGACACGCTGCATCCCTTCCCAAGATATGGAG 2971
QY 854 LeuPheAlaValLeuCysIleGluThrSerHisTyThrValAlaPheValLysTyThrGlyLys 873
DB 2972 TTAATTTGCTGTTCTTGATAGAAACAGCCATATGTTGCTTTTGTGAAGATGGGAAG 3031
QY 874 AspAspSerAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
DB 3032 GACGATTTCTGCTGGCTTCTTTTGACAGCATGCGCGATCGGGATGTTGGTCAAGATGGC 3091
QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyThrLeuLysMetSerLeu 913
DB 3092 TTCAACATTTCTCAAGTCAACCCATGCCAAGAGTAGGAGACTTTTGAAGATGCTCTGT 3151
QY 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCys 933
DB 3152 GAAGACTTGATTTCTTTGGACTCCAGGAATCAAGGCTGTGCAGCAAGACTGCTTTGT 3211
QY 934 AspAlaTyThrMetCysMetTyThrGlnSerProThrMetSerLeuTyLys 949
DB 3212 GATGCATATATGTGATGTACACAGAGTCCCAACAATGAGTTGTACAAA 3259
RESULT 6
ADQ95919
ID ADQ95919 standard; cdna; 3311 BP.
XX
XX ADQ95919;
AC
XX 07-OCT-2004 (first entry)
DT
XX T cell activation associated cdna #49.
```

XX ss; gene; anti-allergic; antiarthritic; antiasthmatic; antidiabetic;
 KW anti-HIV; antimicrobial; antirheumatic; immunosuppressive;
 KW neuroprotective; gene therapy; T cell activation; diagnosis;
 KW autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;
 KW diabetes; allergic disease; infectious disease; AIDS; chronic rejection;
 KW organ; bone-marrow transplant.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 243..3113
 FT /*tag= a

PN WO2004058805-A2.

XX 15-JUL-2004.

XX 25-DEC-2003; 2003WO-JP016715.

XX 26-DEC-2002; 2002JP-00376365.

PR 27-DEC-2002; 2002US-0436473P.

PR 25-APR-2003; 2003JP-00122113.

PR 28-APR-2003; 2003US-0465792P.

PR 21-OCT-2003; 2003JP-00360559.

PR 22-OCT-2003; 2003US-0512846P.

XX (ASAH-) ASahi Kasei PHARMA CORP.

XX Matsuda A, Yoneta S;

XX WPI; 2004-S93134/57.

DR P-PSDB; ADQ95920.

XX Claim 4; SEQ ID NO 97; 2828pp; English.

PT New purified protein involved in T cell activation, useful for
 PT diagnosing, preventing and/or treating acquired immunodeficiency
 PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
 PT and infectious diseases.

XX Sequence 3311 BP; 983 A; 650 C; 773 G; 905 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3311
Score:	4968.50	Matches:	947
Percent Similarity:	99.16%	Conservative:	1
Best Local Similarity:	99.06%	Mismatches:	1
Query Match:	98.70%	Indels:	7
DB:	12	Gaps:	5

US-09-671-687A-3 (1-949) x ADQ95919 (1-3311)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
 DB 243 ATGAGTTTCAGGCTATGAGCCAGAAAGATCACTTCACTCCCTACTGGGAAGCGGATT 302
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 303 TTITACTTCTCTTCAAGATTCAGGTTACAGACAAACAAACAAAGCTCCCTTAA 362
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60

363 GTACCGAAGGAAGTAGGACAGTATATTCAAGATCGTTCTGTGGGCGCATTTCAAGGATT 422
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 DB 423 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 482
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 DB 483 GTTCTCTTTGTTGATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACTT 542
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 543 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAAGCTGTTTAAAAACAGAAAACAGACTAAGT 602
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 603 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTCAGATCTGGGGAGAA 662
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 663 AAATTTCTGGAGTTGTACGCTTTCAGAGGAGCCCTGTTAGCAGAGAGACAGTCTCCGGA 722
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 DB 723 ATATTCTTTGGAGTTGAAATTCCTGGAACAGAGTCTGTCTCAAGTTTCTACATGACGGGTG 782
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 783 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 842
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 843 AAGCTAGAACTCATAGAGATGATGACACTGCACTTGGAAGTGATTAACGAGGCTCTGGG 902
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 903 GACAAATGAGGTGCAACTTCTCTTGGAAATAAACTCCAGAGTTCTTTGAAGTT 962
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 963 GGAGAAACAATAGAACTCGGAACAGTTATATCTGTGATGTTTGGCCAGGAAAGAAAGC 1022
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1023 TTAGGATATTTCTGTGTGGACATGATAACCTATTGGCAACTGGGATGGGAAGATT 1082
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1083 GATGGAGTGCAGCTTTGTTAGTTTGGCGTGTGTTGAAAGTACAAATCTATTGCATCAAT 1142
 QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
 DB 1143 GATATCATCCAGCTTTATCAGAGAGTGTGACGAGGAAGAGGAGCGCTCCCAAACTGCC 1202
 QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
 DB 1203 TTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCAGTCATAATAAACCAAGAGCTACA 1262
 QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
 DB 1263 GGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAAATTTATTTATACCTTAAATGG 1322
 QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
 DB 1323 TCTTCTGTTGACTACAAACCAATCCAAATCAAAAATATCATGATGATGATGATGATGAT 1382
 QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
 DB 1383 GCAGAGAGACCTGCAGAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTACCA 1442
 QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
 DB 1443 CCACTCCAGGCTCTCTCTGTGAACTCACTGACCAACCGAGAACAGATTCCTCTTTTACCA 1502

QY 414 PheSerLeuThrIleValMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
Db 1503 TTCAAGTCTCACCAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTG 1562
QY 434 SerAlaGlnSerValMetGluGluLeuLeuAsnThrAlaProValGlnGlnSerProLeu 453
Db 1563 TCAGCCAGTCTGTAATGGAAGAGCTAAACACTGCACCCGCTCAAGAGAGTCCACCCCTTG 1622
QY 454 AlaMetProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1623 GCCATGCTCTGGGAACCTCACATGCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAG 1682
QY 474 AsnProProPheThrGlyValIleAe-gTTPilleGlyGlnProProGlyLeuAsnGluVal 493
Db 1683 AACCCCTCCTTTCTATGGGGTAAATCCGTTGGATCGGTGAGCCAGCCAGGACTGAATGAAGTG 1742
QY 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db 1743 CTCGCTGGGACTGGAATCGGAAGATAGTGTGCAAGGCTGTACGGATGGAACCTTCAGAGGC 1802
QY 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
Db 1803 ACTCGGTATTTACCTGTGCCCTGAAGAGGCGCTGTTGTGAACCTGAAGAGCTGCAGG 1862
QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 1863 CCTGACTCTAGGTTTCATCTGAGCGGTTTCCAACTCAGATTGAGCGCTGTAACTCT 1922
QY 554 LeuAlaPheGlyClyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
Db 1923 TTAGCATTTGGAGGCTACTTAAGTAGTAGTAGTAGAAGAAATACTCCACCAAAATGGAA 1982
QY 574 LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer 593
Db 1983 AAAGAAGGCTTTGGAGATTAATGATGGGAAGAAAGGCAATCCAGGCTCAATTCAATCT 2042
QY 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2043 TGTACTTAGACTCAACCTTATCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTG 2102
QY 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 633
Db 2103 TTACTTTAGACCCCAAGAAAGACGATGTAGAAATATTATAGTGAAACCCCAAGAGCTACTG 2162
QY 634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
Db 2163 AGACAGAAATTTGTAATCTCTGAGAATATATGGATATGTGTGTCACAAAAATTTATG 2222
QY 654 LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
Db 2223 AAAGTGAAGAAATATCTTGAAGGTGGAGCTGCATCAGATTACCTCTGAAGAAAA 2282
QY 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2283 GATCCTGAGGAATCTTGAATATCTGTTTCATCATATTTAAGGGTAGAACCTTTGCTA 2342
QY 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2343 AAAATAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAA 2402
QY 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuGluTyrPhePheIleAsn 733
Db 2403 AAAAATGAGAAAGTTGGCGTTCACCAATTCAGCAGTTGTAGATGGTCTTTTATCAAC 2462
QY 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
Db 2463 AGTAACCTGAAATTTGAGAGGACCATCATGCTGATTATTTCAGATGCTCTGATTGGA 2522
QY 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2523 AAAGACTTTAACTATTTAAAAAATTTTCTCTGGAATTAATAATAACAGATTTA 2582

QY 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2583 CTTGAAGACACTCCCAAGACAGTCCCGATATGTGGAGGGCTTCAATGATGATGATGATG 2642
QY 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2643 GAATGCTAGCAGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTGTAACCTGC 2702
QY 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2703 AACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCAATAATAACCCAGTGTCACTT 2762
QY 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2763 CCCAAAGACTTACCCGACTGGGACTGGAGACAGCTGCATCCCTTCCAGAAATATGGAG 2822
QY 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2823 TTATTTGCTGTTCTGTCATAGAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAG 2882
QY 874 AspAspSerAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 2883 GACGATTCCTGCTGGCTCTTTTGACAGCATGCCGATCGGATGCTGGTCAAGATGGC 2942
QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu 913
Db 2943 TTCAACATTCCTCAAGTCAACCCATGCCAGAAAGTAGGAGACTTTGAAAGATGCTCTG 3002
QY 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCys 933
Db 3003 GAAGACTGCTATTCCTTGGACTCCAGAGAAATCCAGGCTGTGCACGAGACTGCTTTGT 3062
QY 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3063 GATGCATATATGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3110

RESULT 7
AAA08589
ID AAA08589 standard; DNA; 4527 BP.
AC AAA08589;
XX
XX 19-JUL-2000 (first entry)
XX Human cytoskeleton associated protein 9 (CVSKP-9) coding sequence.
XX Cytoskeleton associated protein; CVSKP-9; cancer; proliferative;
XX autoimmunity; inflammatory; vesicle trafficking; neurological;
XX cardiovascular; cell motility; reproductive; muscle disorder; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 926..3121
XX /*tag= a
XX /product= "CVSKP-9"
XX
XX WO200017355-A2.
XX
XX 30-MAR-2000.
XX
XX 17-SEP-1999; 99WO-US021565.
XX
XX 18-SEP-1998; 98US-0172226P.
XX 27-APR-1999; 99US-0131321P.
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Tang YT, Yue H, Hillman JL, Bandman O, Corley NC;
XX Guegler KJ, Patterson C, Azimzai Y, Baughn MR;
XX WPI; 2000-283582/24.
XX P-PSDB; AAY91954.
XX
XX

3 No
102 (e) date
available

XX Human cytoskeleton associated proteins, used to treat cell proliferative,
 PT autoimmune/inflammatory, vesicle trafficking, neurological, cell
 PT motility, reproductive and muscle disorders.
 XX
 PS Claim 9; Page 107-109; 113pp; English.

XX AAA08581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-
 CC 1 to CYSKP-16) respectively. The sequences can be used to treat and
 CC diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle
 CC trafficking, neurological, cardiovascular, cell motility, reproductive
 CC and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to
 CC CYSKP-16 can be used to treat or prevent disorders associated with
 CC decreased expression or activity of CYSKP (claimed), for example,
 CC atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancers,
 CC autoimmune/anti-inflammatory disorders such as allergies, anemia, asthma,
 CC acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes
 CC mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.
 CC CYSKP antagonists can be used to treat or prevent a disorder associated
 CC with increased expression or activity of CYSKP (claimed)

SQ Sequence 4527 BP; 1351 A; 823 C; 1005 G; 1348 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4527
 Score: 4942.00 Matches: 944
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 5
 Query Match: 98.17% Indels: 5
 DB: 3 Gaps: 4

US-09-671-687A-3 (1-949) x AAA08589 (1-4527)

XX	202	LeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet	221
PT	869	CTCATAGAAAGATGATGACTGCTGCAATGAAAGTACGAGGTCCTGGGGACAAATG	928
PT			
XX	222	GlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGlyGlyThr	241
PS	929	CAGGTGCAACTTCTCTTTGGAATTAACCTCCAGAGTTCTTTGAAGTTGGAGAACAC	988
XX			
CC	242	IleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyr	261
CC	989	ATAGAATCTGGAACAGTTATATTCTGTGATGTTTGGCAGGAAAGAAAGCTTAGGATAT	1048
CC			
CC	262	PheValGlyValAspMetAspAsnProIleGlyAsnThrAspGlyValPheAspGlyVal	281
CC	1049	TTTGTGGTGTGACATGGATAACCTTATGGCACTGGGATGGAAGATTGATGGAGTG	1108
CC			
CC	282	---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIle	300
CC	1109	CAGCTTTGTAGTTTGGTGTGTTGAAAGTACATTTCTATTGCACATCAATGATATCATC	1168
CC			
CC	301	ProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyVal	320
CC	1169	CCAGAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGCCTTTATGTCAAGAGGTGTT	1228
CC			
CC	321	GlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspPro	340
CC	1229	GGGGACAAAGGTTTCATCCAGTCATAATAACCAAGGCTACAGGATCTTACCTCAGACCT	1288
CC			
CC	341	GlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGln	359
CC	1289	GGAAATAGAAACAGATCTGAATTTATACCTTTAAATGGGTCTTCTGTGACTCACA	1348
CC			
CC	360	ProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLys	379
CC	1349	CCACAATCCAATCAAAATACATGTTGATGAAGTTGCAGAGAGCCCTGCACAA	1408
CC			
CC	380	SerLeuThrGluLeuSerThrAspPheAspArgSerSerProProLeuGlnProPro	399
CC	1409	TCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCTCCAGCTCCTCT	1468
CC			
CC	400	ValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMet	419
CC	1469	GTGAACCTCACTGACCCAGAGAGAGATTCACCTCTTTTACCATTCAGTCTCACCAGATG	1528
CC			
CC	420	ProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMet	439
CC	1529	CCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAGCTCTGTAATG	1588
CC			
CC	440	GluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsn	459
CC	1589	GAAGAGCTAAACACTGCACCCGTCGAAGAGAGTCCACCTTGGCCATGCTCTCTGGGAAC	1648
CC			
CC	460	SerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGly	479
CC	1649	TCATGTTGTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTTCTATGGG	1708
CC			
CC	480	ValIleArgThrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeu	499
CC	1709	GTAACTCGTTGGATCGGTGAGCCAGGAGTGAATGAAGTGCCTCGCTGGAGTGAATG	1768
CC			
CC	500	GluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCys	519
CC	1769	GAAGATGAGTGTGAGGCTGACGGATGGAACCTTCAGAGGCACTCGGTATTTCACTGT	1828
CC			
CC	520	AlaLeuLysLysAlaLeuPheValLysSerLysCysArgProAspSerArgPheAla	539
CC	1829	GCCCTGAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTGA	1888
CC			
CC	540	SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyTyr	559
CC	1889	TCATTGCGAGCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTTGGAGGCTAC	1948
CC			
CC	560	LeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIle	579

Db 1949 TTAAGTCAAGTAGAGAGAAATACTCCACAAAATGGAAGAGGCTTGGAGATA 2008
QY MetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThr 599
Db 2009 ATGATTGGGAAGAAGAGGATCCAGGGTCATTACAATTTCTGTACTTAGACTCAACC 2068
QY LeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArgProLysGlu 619
Db 2069 TTATTCCTGCTATTTCCTTTAGTCTGTGTCGACACTGTGTACTTAGACCCAAAGAA 2128
QY LysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsn 639
Db 2129 AGAACGATGTAGAAATATTATAGTGAACCCAGAGCTACTGAGACAGAAATTTGTAAT 2188
QY ProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeu 659
Db 2189 CCTCTGAGAAATATGGAATGTGTGTGCCACAAAATTTATGAACTGAGAAATACTT 2248
QY GluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeu 679
Db 2249 GAAAGGTGGAGGCTCATCAGGATTTACCTCTGAAGAAAAGATCCTGAGGAATCTTG 2308
QY AsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGly 699
Db 2309 AATATTCTGTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGATCAGCAGGT 2368
QY GlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGly 719
Db 2369 CAAAAGGTACAAGATTGTACTTCTATCAANTTTTATGAAAAAATGAGAAAGTTGGC 2428
QY ValProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAla 739
Db 2429 GTTCCCACAATTCAGCAGTTGTTAGAAATGCTTTTATCAACAGTAACCTGAAATTTGCA 2488
QY GluAlaProSerCysIleuIleGlnMetProArgPheGlyLysAspPheLysLeuPhe 759
Db 2489 GAGGCACCATCATGCTGTATTAATTCAGATGCTCGATTTGGAAGAACCTTTAAACTATTT 2548
QY LysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArg 779
Db 2549 AAAAAAATTTTCTCTCTCGAAATTAATAATATACAGATTACTTGAAGACACTCCCGA 2608
QY GlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspPro 799
Db 2609 CAGTGCAGGATATGTGGGGCTTGCAATGTATGAGTGTAGAAATGCTACGAGATCCG 2668
QY AspIleSerAlaGlyLysLysLysGlnPheCysLysThrCysAsnThrGlnValHisLeu 819
Db 2669 GACATCTCAGCTGGAATAATCAAGCAGTTTGTGTAACCTGCAACACTCAAGTCCACCTT 2728
QY HisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAsp 839
Db 2729 CATCCGAAGAGGCTGAATCAATAATAATAACCCAGTGTCACTTCCCAAGACTTACCCGAC 2788
QY TrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCys 859
Db 2789 TGGGACTGGACACCGCTGCATCCCTTGCAGAAATATGAGATTATTTGCTGTCTCTGC 2848
QY IleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeu 879
Db 2849 ATAGAAACAGCCACTATGTGCTTTGTGAAGTATGGGAAGCAGGATTTGCTGCTGCTC 2908
QY PhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnVal 899
Db 2909 TTCTTTTACACAGCATGCCGATCGGGATGGTGGTGCAGAAATGGCTTCAACATTCCTCAAGTC 2968
QY ThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeu 919
Db 2969 ACCCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGAAGACCTGCAATTCCTTG 3028
QY AspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMet 939

Db 3029 GACTCAGGAGAAATCCAGGCTGTGCAGGAAGACTGCTTTGTGATGATATATGTGCATG 3088
QY 940 TyrGlnSerProThrMetSerLeuTyrLys 949
Db 3089 TACCAGATCCAAACAATGAGTTTGTACAAA 3118
RESULT 8
AAH18625
ID AAH18625 standard; cDNA; 2845 BP.
XX AAH18625;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:18842.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isoai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX Claim 8; SEQ ID NO 18842; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX Sequence 2845 BP; 854 A; 557 C; 662 G; 772 T; 0 U; 0 Other;
SQ Alignment Scores: 0 Length: 2845
Pred. No.: 0

Score: 4601.00 Matches: 875
Percent Similarity: 99.32% Conservative: 0
Best Local Similarity: 99.32% Mismatches: 2
Query Match: 91.40% Indels: 4
DB: 4 Gaps: 4
US-09-671-687A-3 (1-949) x AAH18625 (1-2845)
QY 73 LysIleLeuGluGlnProHisAlaValLeuPheValAspGlu---AspValValGluIle 91
Db 2 AAAATTTCTAGAGCAACCTCATGTCAGTCTCTTTGTTGATGAAAGGATGTTGTAGAGATA 61
QY 92 AsnGluLysPheThrCluIleuLeuAlaIleThrAsnGluGluGluArgPheSerLeu 111
Db 62 AATGAAAAGTTTACAGAGTTACTTTTGGCAATTACCAATTTGTGAGAGAGTTTCAGCCCTG 121
QY 112 PheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGlyCysProValLys 131
Db 122 TTTTAAACACAGAAACAGACTAGTAAAGGCTCCAAATAGACGTGGCTGTCTCTGTGAAA 181
QY 132 ValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPheArgGlyProLeu 151
Db 182 GTACAGCTCAGATCTGGGGAAGAAAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTG 241
QY 152 LeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeuGluGlyArg 171
Db 242 TTAGCAGAGGACAGCTCTCCGGAATATTCTTTGGAGTTGAAATGCTGGAAGAGGTCGT 301
QY 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAsp 191
Db 302 GGTCAAGGTTTCACTGACCGGGGTACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGAT 361
QY 192 CysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeu 210
Db 362 TGTGGCGTCTTTGTCATGTGACAGCTAGAACTCATAGAGATGATGACACTGCATTG 421
QY 211 GluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProLeuGluIle 230
Db 422 GAAAGTGATTACGAGGTCCTGGGACACAAATGACAGTCTCGAATCTTCTCTTTGGAAATA 481
QY 231 AsnSerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCys 250
Db 482 AACTCCAGAGTTTCTTTGAGGTTGGAGAAACAAATAGAACTCTGGAACAGTTATATTCTGT 541
QY 251 AspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnPro 270
Db 542 GATGTTTTCCAGGAAAGAAAGCTTAGGATATTTTGGTGTGACATGGATACCCCT 601
QY 271 IleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGlu 289
Db 602 ATTTGCNACTGGGATGGAGATTGATGGAGTGCAGCTTTTGTAGTTTTCGGTGTGTGAA 661
QY 290 SerThrIleLeuLeuHisIleAsnAspIlelleProGluSerValThrGlnGluArgArg 309
Db 662 AGTACAAATTTCTATTGACATCAATGATATCATCCAGAGAGTGTGACGCGAAGAGAGG 721
QY 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsn 329
Db 722 CTCCCAAACTTGCCTTTATGTCAAGAGGTGTGGGGACAAAGGTTTCATCCAGTCAATA 781
QY 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPhe 348
Db 782 AAACCAAGGCTACAGGATCTACTCAGACCTCGGAAATAGAAACAGATCTGAAATATTAT 841
QY 349 TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrp 368
Db 842 TATACCTTAATGGGCTCTCTGTGACTCACAACCACTCCAAATCAAAATAATACATGG 901
QY 369 TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe 388
Db 902 TACATTGATGAAGTTGCAGAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTT 961
QY 389 AspArgSerSerProProLeuGlnProProProValAsnSerLeuThrThrGluAsnArg 408

Db 962 GACCGTTTTCACACACACTCCAGCCTCCTCTCTGTGAACCTCCTACACCCGAGAACAGA 1021
QY 409 PheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHis 428
Db 1022 TTCACACTCTTTTACCATTTCAGTCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCAC 1081
QY 429 SerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGln 448
Db 1082 AGTCCACTTTCTCTGTGAGCCAGCTCTGTATGGAAGAGCTAAACACTGCACCCCTCCAA 1141
QY 449 GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeu 468
Db 1142 GAGAGTCCACCTTGGCCATGCTCTCTGGAACTCACAATGGTCTAGAAAGTGGGCTCATGTG 1201
QY 469 AlaGluValLysGluAsnProProPheTyrGlyValIleArgTyrIleGlyGlnProPro 488
Db 1202 GCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAAATCCGTTGGATCGGTGACGACCA 1261
QY 489 GlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAsp 508
Db 1262 GGACTGAATGAAGTCTGCTGCTGCTGAACTGGAAGATGAGTGTGCAAGGCTGTACGGAT 1321
QY 509 GlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLys 528
Db 1322 GGAACCTTTCAGAGGCACTCGGTATTTCCACTGTGCTGCTGGAAGAGGCGCTGTGTGTGAAA 1381
QY 529 LeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIle 548
Db 1382 CTGAAGAGCTGAGGCTGACTCTAGGTTTGCATCATTTGACGCGGTTTCCATCAGATT 1441
QY 549 GluArgCysAsnSerLeuAlaPheGlyTyrLeuSerGluValValGluGluAsnThr 568
Db 1442 GAGCGCTGTAACCTTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACT 1501
QY 569 ProProLysMetGluLysGluGluIleMetIleGlyLysLysLysGlyIleGln 588
Db 1502 CCACCAAAATGGAAGAAAGGCTTGGAGATATGATTTGGGAAGAGAAAGGATCCAG 1561
QY 589 GlyHisThrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
Db 1562 GGTCAATTACAAATTTCTTACTTAGACTCAACCTTATTTCGCTTATTTCGTTTAGTTCT 1621
QY 609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
Db 1622 GTTCTGGACACTGTGTACTTTGGACCCAAAGAAAGAACGATGTAGAATATTATAGTAA 1681
QY 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
Db 1682 ACCCAAGAGCTACTGAGGACAGAAAATTTGTTAAATCTCTGAGAAATATATGGATATGTGTGT 1741
QY 649 AlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAsaSerGlyPhe 668
Db 1742 GCCACAAAATTTATGAAACTGAGGAAAATACTTTGAAAAGGTGGAGGCTGCATCAGGATTT 1801
QY 669 ThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg 688
Db 1802 ACCTCTGAAGAAAAGATCTGAGGAATCTTGAAATATTCTGTTTCAATATTATTTAAGG 1861
QY 689 ValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyr 708
Db 1862 GTAGAACCTTTGTGTAATAAATAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTAT 1921
QY 709 GlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGlu 728
Db 1922 CAAATTTTATGAAAAAATGAAAAATGGCGTTCGCCAATTCAGCAGTGTGTAGAA 1981
QY 729 TrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGln 748
Db 1982 TGGTCTTTTATCAACAGTAACTTGAAGATTTGCAGAGGACCATCATGTCTGATTATTGAG 2041
QY 749 MetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeu 768

```
Db 2042 ATGCCTCGATTGGAAAGACCTTTAAACTATTTAAAAAATTTTTCCTTCTCTGGAATTA 2101
Qy 769 AenlleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAla 788
Db 2102 AATATAACAGATTTACTTGAAGACACTCCACACAGTCCGGATATGTGGAGGGCTTGCA 2161
Qy 789 MetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGln 808
Db 2162 ATGTATGAGTGTAGAGAATGCTACGACGATCCGGACATCTCAGCTGGAAAAATCAAGCAG 2221
Qy 809 PheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyr 828
Db 2222 TTTTGTAAACCTGCAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATTAATAT 2281
Qy 829 AsnProValSerLeuProLysAspLeuProAspTyrAspTyrArgHisGlyCysIlePro 848
Db 2282 AACCCAGTGTCACTTCCCAAGACTTTACCCGACTGGGACTGGAGACACGCTGCATCCCT 2341
Qy 849 CysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrVallAlaPhe 868
Db 2342 TGCCAGAATATGGAGTATTTGCTGTCTCTGATAGAAACCAAGCCACTATGTGTCTTTT 2401
Qy 869 ValLysTyrGlyLysAspSerAlaTyrLeuPheAspSerMetAlaAspArgAsp 888
Db 2402 GTGAAGTATGGGAAGGACGATTTGCTGGCTGCTCTTTTGACAGCATGGCCGATCGGAT 2461
Qy 889 GlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyr 908
Db 2462 GTGTGTCAGATGGCTTCAACATCTCAAGTCACTCCATGCCAGATGCCAAGCTGTGCA 2521
Qy 909 LeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAla 928
Db 2522 TTGAAGATGTCTCTGAAGACCTGCACTTCTTGGACTCCAGGAGAAATCCAAGGCTGTGCA 2581
Qy 929 ArgArgLeuLeuCysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyr 948
Db 2582 CCAAGACTGCTTTGTGATCATATATGTGATGTACCAAGTCCAAACAATGAGTTTGTAC 2641
Qy 949 Lys 949
Db 2642 AAA 2644

RESULT 9
AAS41617
ID AAS41617 standard; cDNA; 4716 BP.
AC AAS41617;
XX
DT 17-DEC-2001 (first entry)
DE cDNA encoding novel human enzyme polypeptide #833.
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US0001239.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 13-OCT-2000; 2000US-0239937P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
```



```
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246479P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
XX P-PSDB; AAU23747.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
XX Claim 4; SEQ ID NO 843; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC
```

CC encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy. AAS40785-AAS41684 represent cDNA sequences encoding for the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 4716 BP; 1397 A; 848 C; 1036 G; 1431 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	0	Length:	4716
Score:	3876.50	Matches:	736
Percent Similarity:	95.71%	Conservative:	0
Best Local Similarity:	95.71%	Mismatches:	0
Query Match:	77.01%	Indels:	33
DB:	4	Gaps:	4

US-09-671-687A-3 (1-949) x AAS41617 (1-4716)

QY	184	GlnLeuPheGlnCysAspGluAspCysGly	---PheValAlaLeuAspLysLeuGluLeu	202
DB	2	CAGCTTTTTCAGTGTGATGAAGATTGGCGTGT	TTTGTTCATTGGACCACTAGAACTC	61
QY	203	IleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGln		222
DB	62	ATAGAAGATGATGACACTGCATTGGAAAGTGATTACGACGCTCTCGGGACACAAATGCAG		121
QY	223	ValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGlyGlyThrIle		242
DB	122	GTGGAACCTCTCTCTTTGGAAATAAACTCCAGAGTCTTTTGAAG	-----	166
QY	243	GluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPhe		262
DB	166	-----	-----	166
QY	263	ValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPheAspGlyVal	---	281
DB	167	-----GATAACCTTATGGCACTGGGATGGAGATTGTGAGTGAGTGAG		211
QY	282	LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIlePro		301
DB	212	CTTTGTAGTTTTCGGTGTGTGAAAGTACAATTCATTTCACATCAATGATATCATCCA		271
QY	302	GluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGly		321
DB	272	GAGAGTGTGACGCGAGGAAGAGCGCTCCCAAACTTCGCTTTATGTCAAGAGGTGTGGG		331
QY	322	AspLysGlySerSerHisAsnLysProLysValThrGlySerThrSerAspProGly		341
DB	332	GACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACTTCAGACCTCGA		391
QY	342	AsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnPro		360
DB	392	AATGAACACAGATCTGNAATTTATTTATACCTTAATGGGCTCTCTGTGTGATCTACAACA		451
QY	361	GlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLysSer		380
DB	452	CAATCCAAATCAAAAAATACATGCTACATTGATGAGTTGCAGAGACCCCTGCAAAATCT		511
QY	381	LeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlnProProVal		400
DB	512	CTTACAGAGATATCTACAGACTTTTGACCGTTCTTTCCACCACCACCTCCCTCCTGTG		571

QY 401 AnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420
DB 572 AACTCTACTGACCACGAGACAGATTCCACTCTTTTACCAATTCAGTCTCCACCAAGATGCC 631
QY 421 AnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440
DB 632 AATACCAATGGAAAGTATTGGCCACAGTCCACTTTCTGTGACGCCAGGCTCTGTAATGGAA 691
QY 441 GluLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetProProGlyAsnSer 460
DB 692 GAGCTAAACACATGCACCCGCTCCAAAGAGAGTCCACCCCTTGCCCATGCTCTCTGGGAATCA 751
QY 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheThrGlyVal 480
DB 752 CATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCTTCTATGGGGTA 811
QY 481 IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGlu 500
DB 812 ATCCGTTGGATCGGTGAGCCACGAGCTGAATGAAGTGTCTGCTGGACTGGAACTGGAA 871
QY 501 AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAla 520
DB 872 GATGAGTGTGAGGCTGTACGGATGAACTTCAGAGGCACTCGGTATTTCACCTCTGGCC 931
QY 521 LeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSer 540
DB 932 CTGAAGAGGCGCTGTTGTGAACTGAAGAGTGCAGGCGCTGACTCTAGGTTTGCATCA 991
QY 541 LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeu 560
DB 992 TTGCAAGCCGTTTCCAATCAGATTGAGCGCTGTAACCTTTAGCATTTGGAGGCTACTTA 1051
QY 561 SerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMet 580
DB 1052 AGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAGGCTTGGAGATAATG 1111
QY 581 IleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeu 600
DB 1112 ATTGGGAAGAAAGGATCCAGGCTCATTAACATTTCTTTTACTTAGACTCAACCTTA 1171
QY 601 PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys 620
DB 1172 TTCTGCTTATTGCTTTTGTAGTTCTGTTCTGGACACTGTGTACTAGACCACCAAGAAAG 1231
QY 621 AsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640
DB 1232 AACGATGTAGAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATTTGTAATCCT 1291
QY 641 LeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660
DB 1292 CTGAGAATATATGATATGTGTGCCACCAAAATTTATGAACCTGAGGAAATTTACTTGAA 1351
QY 661 LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn 680
DB 1352 AAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGACTCTGAGGAATTTCTGAAT 1411
QY 681 IleLeuPheHisIleLeuArgValGluProLeuLysIleArgSerAlaGlyGln 700
DB 1412 ATTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGATCAGAGGTCAA 1471
QY 701 LysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyVal 720
DB 1472 AAGGTACAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATGNAAGTTGGCGTT 1531
QY 721 ProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAlaGlu 740
DB 1532 CCCACAATTCAGCAGTGTGTAGAAATGGTCTTTTATCAACAGTAACTGAAATTTGCAGAG 1591
QY 741 AlaProSerCysIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760
DB 1592 GCACCATCATGCTGTGATTTATTCAGATGCGCTCGATTTGGAAAAAGACTTTAAACTATTAA 1651
QY 761 LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780

DB 1652 AAAATTTTCTCTCTCGAAATTAACACAGATTATCTTGAAGACACTCCACAGACAG 1711
QY 781 CysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspProAsp 800
DB 1712 TGCCGGATATGGAGGGCTTGCATGTATGAGTGTAGAGATGCTACGAYGATCCGGAC 1771
QY 801 IleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHis 820
DB 1772 ATCTCAGCTGGAATAATCAAGCAGTTTTGTAAAACTGCAACACTCAAGTCCACCTTCAT 1831
QY 821 ProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrp 840
DB 1832 CCGAAGAGCTGTAATCATAATATAACCCAGTGTCTCTCCCAAGACTTACCCGACTGG 1891
QY 841 AspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860
DB 1892 GACTGAGACACGGCTGTCATCTCTGCCAGAAATATGAGATTATTTGCTGTCTCTGCATA 1951
QY 861 GluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPhe 880
DB 1952 GAAACAAGCCACTATCTTCTTTTGTGAAGTATGGGAAGGACGATCTCGCTGCTCTTC 2011
QY 881 PheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
DB 2012 TTTGACAGCATGGCCGATCGGATGGTGTGTCAGAAATGGCTTCAACATTCTCTCAAGTCA 2071
QY 901 ProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAsp 920
DB 2072 CCATGCCCAAGAGTAGAGAGTACTTGAAGATGTCTCTGGAAGACCTGCAATTCCTTGAC 2131
QY 921 SerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyrMetCysMetTyr 940
DB 2132 TCCAGAGAGAAATCCAAAGCTGTGCAAGAGACTGCTTTGTGATGCATATATGTCATGTAC 2191
QY 941 GlnSerProThrMetSerLeuTyrLys 949
DB 2192 CAGAGTCCCAACATGAGTTGTACAAA 2218

RESULT 10
AAI58410
ID AAI58410 standard; cDNA; 2523 BP.
XX
AC AAI58410;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 613.
XX
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN W0200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM39254.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 613; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,22e-312 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 4 Gaps: 2

US-09-671-687A-3 (1-949) x AAI58410 (1-2523)

QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
DB 45 GGAAATAAAGCTTTTATATTTTGGTGGACATGATGATACCTTATGGCAACTGG 104
QY 275 AspGlyArgPheAspGlyVal----LeuCysSerPheAlaCysValGluSerThrIleLeu 293
DB 105 GATGGAAGATTGATGGAGTGCAGCTTTGTAGTTTTCGGTGTGTGAAAGTACAAATCTA 164
QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeu 313
DB 165 TTGCACATCAATGATATATATCCAGAGAGTGTACCGAGAAAGGAGGCCCTCCCAACTT 224
QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333
DB 225 GCCTTTATGTCAAGAGGTGTGGGCAAAAGGTTCATCCAGTCATATATTAACCAAGGCT 284
QY 334 ThrGlySerThrSerAspProGlyAsnArg--ArgSerGluLeuPheTyrThrLeuAsn 352
DB 285 ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAT 344
QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu 372
DB 345 GGGCTCTCTGTTGACTCACAAACACAAATCCAAATCAAAATAACATGATGATGAA 404
QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392
DB 405 GTTGCAGAGAGACCCCTGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTCTTCA 464
QY 393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412

DB 465 CCACCACCTCCAGCCTCCTCCTGTGAACTCACTGACCACCGAGAACAGATTCCACTCTTTA 524
QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
DB 525 CCATTGAGTCTCACCAGATGCCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTTCT 584
QY 433 LeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProPro 452
DB 585 CTGTCCAGCCAGTCTGTATGGAAGAGCTAAACACTGCACCCGTCGAAGAGAGTCCACCC 644
QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluVallys 472
DB 645 TTGCCCATGCTCCTCGGAACCTCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAG 704
QY 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
DB 705 GAGAACCTCTCTTCTATGGGTAATCCGTGGATCGTCAAGCCACAGGACTGAATGAA 764
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTCTCGCTCGACTGGAACTGGGAAGATGAGTGTGCAGGCTGTACGGATGGAACTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
DB 825 GGCACCTCGGATTTTCACTGTGCCCTGAAGAGCGCTGTTTGGAAACTGAAGAGCTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGCGCTGACTTAGTGTTCATCATTCAGCCGGTTTCCAATCAGATTGAGCGCTGTAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTGGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACTCCACCACAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGAACCATCCAGGGTCAATTACAAAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTACTTAGACTCACTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTACTTAGACCCCAAGAAAGAACGATAGTAAATATATAGTAGAAACCCAGAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACAGAAAATTGTAATCTCTCGAAGATATATGATATGTGTGTGCCACAAAATTT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAAACTGAGGAAATATCTTGAATATTTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTG 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCTCAGGAAATCTTGAATATTTCTGTTTCATCATATTTTAAAGGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAAAATAAGATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnLeuLeuThrSerPheIle 732
DB 1425 GAAAAAATAGAGAAAGTGGCGTTCCCAATTCAGCAGTTGTTAGATGTTGCTTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleLeuGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTTGCAGAGGCACCATCATGCTGATTATTCAGATGCTTCGATTT 1544
QY 753 GlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GGAAAGACITTTAAACTATTTAAAAAATTTTCTCTCTGGAATTTTAAATATAACAGAT 1604

QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
Db 1605 TTACTTTGAAGACACTCCAGACAGTCCCGATATGTGGAGGGCTTGCAATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleGlnPheCysLysThr 812
Db 1665 AGAAGATGCTACAGACATCCGACATCTCAGCTGGAAATCAATCAAGCAGTTTGTAAACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
Db 1725 TGCAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATATAATAACCCAGTGTC 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
Db 1785 CTTCCCAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCCCAATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
Db 1845 GAGTTATTTGCTGTTCTCTGCATAGAAACAAGCCATGTATGCTTTTGTGAAGTATGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
Db 1905 AAGGACGATTCGCTGGCTCTCTTTGACAGCATGGCCGATCGGATGGTGTGAGAT 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
Db 1965 GGCTTCAACATTCCTCAAGTCAACCCATGCCAGAGAGTAGGAGTACTTTGAAGATGCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
Db 2025 CTGGAGACCTGTCATTCCTTTGGACTCCAGGAGAAATCAAGGCTGTGCAACGACTGCTT 2084
QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2085 TGTGATCATATATGTCATGTACAGATCCAGAGTCCCAACATGATGTTGTACAA 2135

RESULT 11

ADQ98620
ID ADQ98620 standard; cDNA; 2523 BP.

XX AC ADQ98620;

XX DT 23-SEP-2004 (first entry)

XX DE DNA encoding human GPCR-like protein seqid 290.

XX KW ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
XX KW antidiabetic; GPCR-like protein; ophthalmic disorder;
XX KW neurological disorder; immunological disorder; nephritic disorder;
XX KW hormonal dysfunction; cancer; atherosclerosis; diabetes;
XX KW molecular weight marker; food supplement; human; ss.

XX OS Homo sapiens.

XX PN US6569662-B1.

XX PD 27-MAY-2003.

XX PF 19-JUL-2000; 2000US-00620312.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Zhou P, Drmanac RT;

XX DR WPI; 2001-442255/47.

XX PT New G-protein-coupled receptor-like polypeptides and polynucleotides,
XX PT useful for treating diseases of ophthalmic, neurological, immunological
XX PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis

PT and diabetes.

XX Example 2; SEQ ID NO 290; 92pp; English.

XX CC The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translation or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.

SQ Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-22e-312 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 5 Gaps: 2

US-09-671-687A-3 (1-949) x ADQ98620 (1-2523)

QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274

Db 45 GGAATAATAAGCTTTTATATTTTGTGTGTGACATGATGATCAACCTATTGGCAACTGG 104

QY 275 AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu 293

Db 105 GATGGAAGATTTGATGGAGTGCAGCTTTGTAGTTTTCGTTGTTGAAGTACAAATCTTA 164

QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProLysLeu 313

Db 165 TTGCACATCAATGATATATATCCAGAGAGTGTGACGAGGAGGAGGAGGCTCCCAACTT 224

QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333

Db 225 GCCTTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCAATATAAACCAAGGCT 284

QY 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352

Db 285 ACAGGATCTACCTCAGACCTCGGAAATAGAAACAGATCTGAATATATTTATACCTTAAAT 344

QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu 372

Db 345 GGGTCTCTGTGACTCACACCAACCAATCCAAATCAAAAATACATGATGATGAA 404

QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392

Db 405 GTTGCAAGAGACCTCGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTCTTCA 464

QY 393 ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerIleu 412

Db 465 CCACCACTCCAGCCTCTCTCTGGAACCTCACTGACCAACCCAGAACAGATTCCTCTTTA 524

QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432

Db 525 CCATTGATCTCACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCT 584

QY 433 LeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProPro 452

Db 585 CTGTCAGCCAGTCTGTAATGGAAGAGCTAAACACTGCACCCCGTCCAAGAGAGTCCACCC 644

QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472

Db 645 TTGGCCATGCTCTCTGGAACTCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAG 704

QY 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
DB 705 GAGAACCCCTCTTCTATGGGGTAATCCGGTTGGATCGGTGAGCCACGAGACTGAATGAA 764
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTGCTCGCTGGACTGGAACCTGGAAGATGAGTGTGACGGCTGTACGGATGGAACCTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysLysLysLysLysLysLysLysLys 532
DB 825 GGCACCTCGGTATTTTCACTTGGCCCTGAAGAGCGCTGTTTGTGAAACTGAAGAGCTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGCCCTGACTCTAGGTTTGCATTCATTCAGCCGGTTCCTCAATCAGATTGAGCGCTGTAA 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluLeuAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTGGAGCTTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGGCTGGAGATAATGATGGGAAGAGAGGATCCAGGGTCAATTACAAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTACTTACACTCAACCTTATTTCTGCTTATTTTGTCTTTTGTCTTGTCTGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTACTTAGACCCCAAGAAAGAACGATGATGATATTTATATGTAAGCCCAAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACAGAAATCTTAACTCTCTGAGATATATGATATGTTGTGTCACAAAAT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAAACTCAGGAAAAATCTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGA 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCCTGAGGAATCTTGAATATCTGTTTCATCATATTTAAGGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAATAATAGATCACAGGCTCAAAAGGTACAGATTGTACTTCTATCAAAATTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCGTTCCCAATTCAGCAGTTGTTAGATGGTCTTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACTCGAAATTTGAGAGGACCCATCATCTGTGATTATTACAGATCGCTCGATT 1544
QY 753 GlyLysAspPheLysLeuPheLysLysIlePheProSerLeuLeuLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAACTATTTAAAAAAATTTTCTCTCTGGAATTAATAATAACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTGAAGACACTCCAGACAGTGGCGATATGTGGGGCTTGCATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAGAATGCTACGACGATCGGACATCTCAGCTGGAAAAATCAAGCAGATTTTGTAAACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCAATCGAAGAGGCTGAATCATAAATATATACCCAGTGTCA 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852

DB 1785 CTTCCTCCAAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTATTTCTGCTCTCTCGATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGGACGATTTCTGCTGGCTCTTCTTTGACAGCATGGCCGATCGGATGGTGGTCAGAA 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGluTyrLeuLysMetSer 912
DB 1965 GGCTTCAACATTTCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTGAAGATGCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
DB 2025 CTGGAAGACCTGCATTTCTTGGACTCCAGGAGATCCAAAGCTGTGCACAGAACTGCTT 2084
QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 2085 TGTGATGCATATATGTGCATGTACAGAGTCCAAATGAGTTTGTACAAA 2135
RESULT 12
ADB48380
ID ADB48380 standard; cDNA; 2523 BP.
XX
AC ADB48380;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human cDNA SEQ ID NO 290.
XX
KW ss; cancer; neurodegenerative disease; human.
XX
OS Homo sapiens.
XX
PN US2003104529-A1.
XX
PD 05-JUN-2003.
XX
PF 04-JAN-2002; 2002US-00037270.
XX
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 19-JUL-2000; 2000US-00620312.
XX
PA (ZHOU/) ZHOU P.
PA (TANG/) TANG Y T.
PA (LIUC/) LIU C.
PA (ASUN/) ASUNDI V.
PA (DRMA/) DRMANAC R T.
XX
PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
XX
DR WPI; 2003-678194/64.
XX
PT New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
XX
PS Claim 1; SEQ ID NO 290; 99pp; English.
XX
CC The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=20030104529.
XX
SQ Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;

late

Alignment Scores:

Pred. No.:	1-22e-312	Length:	2523
Score:	3676.00	Matches:	692
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	3
Query Match:	73.02%	Indels:	2
DB:	9	Gaps:	2
US-09-671-687A-3 (1-949) x ADB48380 (1-2523)			
QY	255	GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp	274
DB	45	GGAAATAAAGCTTTTATATTTTGGTGGACATGGATGAACCTATTGGCAACTGG	104
QY	275	AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu	293
DB	105	GATGGAAGATTGATGGAGTGCAGCTTTGTAGTTTTGGCGTGTGTGAAAGTACAATTCTA	164
QY	294	LeuHisIleAsnAspIlelleProGluSerValThrGlnGluArgArgProProLysLeu	313
DB	165	TTGCACATCAATGATATCATCCAGAGAGTGTGACGCGAGAAAGGAGGCTCCCAACTT	224
QY	314	AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla	333
DB	225	GCTTTTATGTCAAGAGGTGTGGGACAAAGGTTCATCCAGTCATATAAACAAGGCT	284
QY	334	ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn	352
DB	285	ACAGGATCTACCTCAGACCTCGAATATAGAACAGATCTGAATATTTTATACCTTAAT	344
QY	353	GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu	372
DB	345	GGGTCTCTCTGACTCACAAACCAATCCAAATCAAAATAACATGTTACATTGATGAA	404
QY	373	ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer	392
DB	405	GTTGCAGAGACCTCGAAATCTCTTACAGAGATATCTACAGCTTTGACCGCTCTTCA	464
QY	393	ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu	412
DB	465	CCACCACTCCAGCTCTCTCTGTAACCTACTGACCAACCAGAACAGATTCACACTCTT	524
QY	413	ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer	432
DB	525	CCATTGAGTCTCACCAGATGCCAATACCAATGGGAAGTATTGGCCACAGTCCACTTCT	584
QY	433	LeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProPro	452
DB	585	CTGTCAGCCAGTCTGTAAATGGGAAGAGCTAAACACTGCACCCCTCCAAAGAGAGTCC	644
QY	453	LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys	472
DB	645	TTGGCCATGCTCTCGGAACTCACATGGTGTAGAAGTGGGCTCATTTGGCTGAAGTTAAG	704
QY	473	GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu	492
DB	705	GAGAACCTCTTCTATGGGTAAATCCGTGGATCGGTGAGCCACAGGACTGAATGAA	764
QY	493	ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg	512
DB	765	GTGCTCGCTGGAATGGAATGGAATGAGTGTGCGAGGCTGTACGGATGGAACCTTCAGA	824
QY	513	GlyThrArgTyrPheThrCysAlaLeuLysLeuAlaLeuPheValLysLeuLysSerCys	532
DB	825	GGCACTCGGTATTTTCACTTGTGCCCTGAAGAAGCGCTGTTTGTGAACACTGAAGAGCTGC	884
QY	533	ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn	552
DB	885	AGGCTGACTCTAGGTTTGCATTTGCAGCCGGTTCCCAATCAGATTGAGCGCTGTAA	944
QY	553	SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrProProLysMet	572
DB	945	TCCTTTAGCATTTGGAGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATG	1004

QY	573	GluLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsn	592
DB	1005	GAAAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGGCAATCCAGGGTCAATCAAT	1064
QY	593	SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr	612
DB	1065	TCCTGTACTTAGACTCAACCTTTATTCTGCTTTATTTGCTTTTGTCTCTGGACACT	1124
QY	613	ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu	632
DB	1125	GTGTTACTTAGACCCAAAGAAAGAACGATGTAGATATTTATGTAGAACCCAGAGCTA	1184
QY	633	LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle	652
DB	1185	CTGAGGACAGAAATTTGTAATCTCTCAGAAATATATGGATATGTGTGCCACAAAATT	1244
QY	653	MetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu	672
DB	1245	ATGAAACTGAGGAAAAATCTTGAAAAGGTGGAGCTGCATCAGGATTTTACCTCTGAAGAA	1304
QY	673	LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu	692
DB	1305	AAAGATCCTGAGGAATTTCTTGAATATTTCTGTTTCATCATATTTTAAAGGTAGAACCTTTG	1364
QY	693	LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet	712
DB	1365	CTAAAAATAAGATCAGCAGCTCAAAAGGTACAAAGATTTGTACTTCTATCAAAATTTATG	1424
QY	713	GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIle	732
DB	1425	GAAAAAATGAGAAAGTTGGCGTTCCACAAATTCAGCAGTTGTAGAAATGCTTTTATC	1484
QY	733	AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe	752
DB	1485	ACAGTAACCTGAAATTTGCGAGGACCACATCATGTCTGATTTTACAGATGCCTCGATT	1544
QY	753	GlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAsp	772
DB	1545	GGAAAGACTTTAAACTATTTAAAAAAATTTTCTCTCTGGAATTTAAATATAACAGAT	1604
QY	773	LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys	792
DB	1605	TTACTTTGAAGACACTCCACAGAGTCCGATATGTGGAGGCTTGCATATGATCAGTGT	1664
QY	793	ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr	812
DB	1665	AGAGAATGCTTACGACCATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACC	1724
QY	813	CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer	832
DB	1725	TGCAACACTCAAGTCCACCTTCATCCGAAAGGCTGAATCATTAANTATATACCCAGTGCA	1784
QY	833	LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet	852
DB	1785	CTTCCCAAGACTTACCCGACTGGACTGGAGACACGCTGCATCCCTTGGCCAGAAATG	1844
QY	853	GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly	872
DB	1845	GAGTTATTGCTGCTCTCTGCATAGAAACAAGCCACTATGTGCTTTTGTGAAGTATGGG	1904
QY	873	LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn	892
DB	1905	AAGACGATTTCTGCTGGCTCTTCTTTTGACAGCATGGCCCATCGGGATGGTGGTCAAGAT	1964
QY	893	GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer	912
DB	1965	GGCTTCAACATCTCCTCAAGTCAACCCATGCCCCAGAGTAGGAGATGCTTGAAGATGCT	2024
QY	913	LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu	932
DB	2025	CTGGAAGACCTGCTATCTTTGGACTCCAGGAGAAATCCAAGGCTGTGACGAAAGCTGCTT	2084

QY	933	CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys	949
Db	2085	TGTGATGATATATGTCATGACAGAGTCCAAATGAGTTGTACAAA	2135
RESULT 13			
AAI60196			
ID	AAI60196	standard; cDNA; 4286 BP.	
AC	AAI60196;		
XX	22-OCT-2001	(first entry)	
XX	Human polynucleotide	SEQ ID NO 4185.	
DE	Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia; ss.		
OS	Homo sapiens.		
XX	WO200153312-A1.		
PN	26-JUL-2001.		
XX	26-DEC-2000; 2000WO-US034263.		
XX	23-DEC-1999; 99US-00471275.		
PR	21-JAN-2000; 2000US-00488725.		
PR	25-APR-2000; 2000US-00552317.		
PR	20-JUN-2000; 2000US-00598042.		
PR	19-JUL-2000; 2000US-00620312.		
PR	03-AUG-2000; 2000US-00653450.		
PR	14-SEP-2000; 2000US-00662191.		
PR	19-OCT-2000; 2000US-00693036.		
PR	29-NOV-2000; 2000US-00727344.		
XX	(HYSE-) HYSEQ INC.		
PA	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;		
PI	Zhou P, Goodrich R, Drmanac RT;		
XX	WPI; 2001-442253/47.		
DR	P-PSDB; AAM41040.		
XX	Novel nucleic acids and polypeptides, useful for treating disorders such		
PT	as central nervous system injuries.		
XX	Claim 1; SEQ ID NO 4185; 10078pp; English.		
PS	The invention relates to human nucleic acids (AAI57798-AAI61369) and the		
CC	encoded polypeptides (AAI38642-AAI2213) with neotropic,		
CC	immunosuppressant and cytoskeletal activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression.		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders. Note: The sequence data for this patent did not form		
CC	part of the printed specification		
XX	Sequence 4286 BP; 1279 A; 779 C; 916 G; 1312 T; 0 U; 0 Other;		
SQ	Alignment Scores:		
XX	Pred. No.: 7.57e-299 Length: 4286		

Score: 3523.00 Matches: 685
Percent Similarity: 97.31% Conservative: 2
Best Local Similarity: 97.03% Mismatches: 8
Query Match: 69.98% Indels: 11
DB: 4 Gaps: 4

US-09-671-687A-3 (1-949) x AAI60196 (1-4286)

QY	255	GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp	274
Db	44	GGAAAAATAAAGCTTTTATATTTTGTGGTGGACATGGATAACCCCTATTGGCAACTGG	103
QY	275	AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu	293
Db	104	GATGGGAAGATTGATGGAGTGCAGCTTTGTAGTTTTCGTGTGTGTGAAGATCAATTTCTA	163
QY	294	LeuHisIleAsnAspIleLeuProGluSerValThrGlnGluArgArgProProLysLeu	313
Db	164	TTGCACATCAATGATATCATCCAGAGAGTGTGACGCGAGGAAAGAGGCGCTCCCAAACTT	223
QY	314	AlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAla	333
Db	224	GCCTTTATGTCAAGAGGTGTGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGCT	283
QY	334	ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn	352
Db	284	ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAAATTTATTATACCTTAAT	343
QY	353	GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu	372
Db	344	GGGTCTCTGTGACTCACAAACCAATCCAAATCAAAAATAATACATGATGATCATGATGAA	403
QY	373	ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspSerSer	392
Db	404	GTTGCAGAGACCCCTGCANAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCA	463
QY	393	ProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu	412
Db	464	CCACCACCTCCAGGCTCTCTCTGTGAACTCACTGACCACCGAGAACAGATTCACCTCTTTA	523
QY	413	ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProIleSer	432
Db	524	CCATTGAGTCTCACCAGATGCCCAATACCAATGGAAGATATTGGCCACAGTCCACTTTCT	583
QY	433	LeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGlnSerProPro	452
Db	584	CTGTGAGCCCGTCTGTATGGAAGAGCTAAACACTGCACCCCTCCAGAGAGTCCACCC	643
QY	453	LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys	472
Db	644	TTGGCCATGCCTCCTGGAACTCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAG	703
QY	473	GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLysAsnGlu	492
Db	704	GAGAACCTCTCTTCTATGGGGTAATCCGTGTGATCGGTGAGTCCAGCCAGGAGTCAATGAA	763
QY	493	ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCys---ThrAspGlyThrPheAr	512
Db	764	GTGCTCGTGGATGGAACTGGAAAGATGAGTGTGAGGCTGGTACGATGGAACCTTCAA	823
QY	512	g---GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSer	531
Db	824	GAAGGCACCTCGGTATTTCACCTGTGCCCTGAAGAAGCGCTGTTTGTGTAAGTGAAGAGC	883
QY	532	CysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCys	551
Db	884	TGCAGGCTGACTCTAGTTTGCATCATTTGACGCGGTTTCCCAATCAGATTGAGCGGTGT	943
QY	552	AsnSerLeuAla---PheGlyGlyTyrLeuSerGluValValGluGluAsnThrPro---	569
Db	944	AATCTTTAGCCATTTGGGAGGCTTACTTAAGTGAAGTAGTGGAGGAAATACTCCACACC	1003
QY	570	ProLysMetGluLysGluGlyLeuGluIleMetIle-GlyLysLysLysGlyIleGlnGln	589

Db 1004 CAAAAATCGGAAAAAGAGGCTTGAGATATATGATGGGAGAGAAAGGCATCCAGGG 1063
Qy yHisTyrAenSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerVa 609
Db TCATTACAAATCTTGTTACTTAGACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGT 1123
Qy lLeuAspThrValLeuLeuArgProLysGluLysAenAspValGluTyrTyrSerGluTh 629
Db TCTGCACACTGTGTTACTTAGACCCAAAGAAAGAACGATGTAGAAATATATAGTGAAC 1183
Qy rGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAl 649
Db CCAAGAGCTACTGAGGACAGAAATTTGTAATCCTCTGAGAATATATGGATATGTGTGTC 1243
Qy aThrIysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheTh 669
Db CACAAAAATATGAAACTGAGGAAATACTTGAAAGGTGGAGGCTGCATCAGGATTTAC 1303
Qy rSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgVa 689
Db CTCTGAAGAAAGATCCTGAGNAATCTTGAATATCTGTTTCATCATATTTTAAGGCT 1363
Qy lGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGl 709
Db AGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCA 1423
Qy nIlePheMetGluLysAenGluLysValGlyValProThrIleGlnGlnLeuLeuGluTr 729
Db AATTTTATGGAAGAAAAATGAGAAAGTTGGCGTCCCAATTCAGCAGTTGTTAGAAATG 1483
Qy pSerPheIleAenSerAenLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMe 749
Db GTCCTTTTATCAAGTAACTGAAATTTGCAGAGGCACCATCATGCTGATTTATTCAGAT 1543
Qy tProArgPheGlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAs 769
Db GCCTCGATTTGAAAGAGCTTTTAAACTATTTAAAAAAATTTTCTCTCTGGAATTTAA 1603
Qy nIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMe 789
Db TATAACAGATTTTACTTGAAGACACTCCAGACAGTCCCGGATATGCGAGGCGCTTCAAT 1663
Qy tTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPh 809
Db GTATGAGTGAGAAATGCTACACATCCGACATCTCAGCTGGAAAAATCAAGCAGTT 1723
Qy eCysLysThrCysAenThrGlnValHisLeuHisProLysArgLeuAenHisLysTyrAs 829
Db TTGTAAAACTCGAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATAAATATAA 1783
Qy nProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCy 849
Db CCCAGTGTCACTTCCCAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTG 1843
Qy sGlnAenMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheVa 869
Db CCAGAAATATGGAGTATTTGCTGTTCTCTGCATAGAAACAAAGCCACTATGTTCTTTGT 1903
Qy lLysTyrGlyLysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGl 889
Db GAAGTATGGAGGACGATCTCGCTGGCTCTCTTTTGACAGATGGCCGATCGGGATGG 1963
Qy yGlyGlnAenGlyPheAenIleProGlnValThrProCysProGluValGlyGluTyrLe 909
Db TGGTCAGATGGCTTCAACATTCCTCAAGTCACCCATGCCATGCCAGATAGGAGTACTT 2023
Qy uLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaAr 929
Db GAAGATGTCTCTGGAAGAGCTGATCTCTTGGACTCCAGAGAAATCAAGGCTGTGCACG 2083
Qy gArgLeuLeuCysAspAla-TyrMetCys-MetTyr-GlnSerProThrMetSerLeuTy 948

Db 2084 AGACTGCTTTGTGATGCCATATATGTCATGTACCCAGAGTCCCAACAATGAGTTTGT 2143
Qy 948 rLys 949
Db 2144 CAAA 2147
RESULT 14
AAS41081
ID AAS41081 standard; cDNA; 2488 BP.
XX AAS41081;
AC AAS41081;
XX 17-DEC-2001 (first entry)
XX cDNA encoding novel human enzyme polypeptide #297.
XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX Homo sapiens.
OS
XX WO200155301-A2.
PN
XX 02-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US001239.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465566/50.

P-PSDB; AAU23211.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.

Claim 4; SEQ ID NO 307; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy. AAS40785-AAS41684 represent cDNA sequences encoding for the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 2488 BP; 708 A; 520 C; 601 G; 656 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	3,86e-282	Length:	2488
Score:	3329.00	Matches:	668
Percent Similarity:	70.09%	Conservative:	0
Best Local Similarity:	70.09%	Mismatches:	3
Query Match:	66.13%	Indels:	283
DB:	4	Gaps:	6

US-09-671-687A-3 (1-949) x AAS41081 (1-2488)

Qy	1	MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgile	20
Db	246	ATGATGTCAGCTTATGAGCCCAAGAAAAGTCTACTTCACCTACTGGAGAGCGGATT	305
Qy	21	PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrLysLeuLeuLys	40
Db	306	TTTTTACTTGCTTCTTCAAGAAATGCAGCGTTACACACAAACACAAAGTCCTTAAA	365

QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 366 GTACCAAGGAGTAGTACACAGTATATTACAGATCGTTCTGTGGGGCATTTCAARGATT 425
QY 61 ProSerAlaLysGlyLysLysAenGlnIleGlyLeuLysIleLeuGlnProHisAla 80
DB 426 CCTTCTGCAAAAGGCAAGAAATACAGATTGGATTAAATAATTCTAGACAACTCATGCA 485
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
DB 486 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 545
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 546 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCAGCCTGTTTAAAAACACAGACAGACTAAGT 605
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 606 AAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 665
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 666 AATTTCTTGGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGGACAGTCTCCGGA 725
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 726 ATATTTCTTTGGAGTTGAAATGCTGGAAGAGGTCGTGGTCAAGGTTTCTACTGACGGGGTG 785
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 786 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGATTTGGAC 845
QY 199 LysLeuLeuLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 846 AAGCTAGAAGTCTAGAAGATGATGACACTGCAATGGAAAGTATTACGCAGGTCTCTGGG 905
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 906 GACACAATGCAGTGCAGATCTCTCTCTTGGAAATAAATCCAGAGTTTCTTTGAAG--- 962
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 962 ----- 962
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 963 -----GATAACCTTATTGGCAACTGGGATGGAAGATT 995
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 996 GATGGAGTGCAGCTTTGTAGTTTGGTGTGTTGAAAGTACAATTTCTATTGGACATCAAT 1055
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
DB 1056 GATATCATCCACAGAGTGTGACGACGAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1115
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1116 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC 1175
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1176 TCAGACCTGGAATAGAAACACATCTGAATATTTTATACCTTAATGGGTCTTCGTT 1235
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1236 GACTCACAACCAATCCAAATCAAAAATACATGTTGATGAAGTTGCAGAGAC 1295
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1296 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCTCCAG 1355

QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1356 CCTCTCTCTGTGAACCTCACTGACCAACGAGAACAGATTCCACTCTTTACCATTCAGTCTC 1415
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1416 ACCAAGATGCCCATACCAATGGAAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAG 1475
QY 437 SerValMetGluGluLeuLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetPro 456
DB 1476 TCTGTAAATGGAGAGCTAAACACTGCACCGCTCAAAGAGAGTCCACCTTTGGCCATCCCT 1535
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1536 CCTGGGAACCTCATGCTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT 1595
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1596 TTTCTATGGGGTAATCCGTTGGATCGGTACGCCACAGGACTGAATGAAGTGTCTCGCTGGA 1655
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1656 CTGGAACCTGGAAGATCAGTGTGCAGGCTGTACGGATGGAACCTTCAGAGGCACCTCGGTAT 1715
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1716 TTCACCTGTGCTGCTGAAGAAGGCGCTGTTGTGTAACCTGAAGAGCTGCAGGCTGACTCT 1775
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1776 AGGTTTGGCATCATTTGCAGCGCGTTTCCAATCAGATTGAGCGCTGTAACCTCTTAGCATTT 1835
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 1836 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACTT----- 1878
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1878 ----- 1878
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 1878 ----- 1878
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 1878 ----- 1878
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 1878 ----- 1878
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 1878 ----- 1878
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 1878 ----- 1878
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 1878 ----- 1878
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 1878 ----- 1878
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
DB 1878 ----- 1878
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776

Db 1878 ----- 1878
 QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
 Db 1878 ----- 1878
 QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
 Db 1878 ----- 1878
 QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
 Db 1879 -----CATCCGAGAGGCTGATCATATATAACCCAGTGTCACTTCCCAAGAC 1929
 QY 837 LeuProAspTTPAspTTPArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
 Db 1930 TTACCCGACTGGGACTGGAGACACGCGCTGCATCCCTTGCAGAAATATGGAGTTATTGCT 1989
 QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 Db 1990 GTTCTCTGCATAGAAACAGCCCACTATGTGCTTTTGTGAAGTATGGGAGGACGATTCT 2049
 QY 877 AlaTTPLeuPhePheAspSerMetAlaAspAtgAspGlyGlnAsnGlyPheAsnIle 896
 Db 2050 GCTGCTCTTCTTTGACAGCATGGCGGATCGGATGGTGGTGCAGATGCTTCAACATT 2109
 QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 Db 2110 CCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAGACCTG 2169
 QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 Db 2170 CATTCCTTGGACTCCAGGAGATCCCAAGGCTGTGCACGAGACTGCTTGTGATGCATAT 2229
 QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 2230 ATGTGCATGTACCAGATGCCAACATGAGTTGTACAAA 2268

RESULT 15

AAH18478

ID AAH18478 standard; cDNA; 2569 BP.

XX AC AAH18478;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:18586.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PI Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 CDNAs.

PS Claim 8; SEQ ID NO 18586; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification, where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 2569 BP; 788 A; 470 C; 562 G; 749 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.6e-255 Length: 2569
 Score: 3022.50 Matches: 587
 Percent Similarity: 98.66% Conservative: 2
 Best Local Similarity: 98.32% Mismatches: 3
 Query Match: 60.04% Indels: 5
 DB: 4 Gaps: 4

US-09-671-687A-3 (1-949) x AAH18478 (1-2569)

QY 88 ValValGluLeuAsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGluGlu 107
 Db 1 GTTGTAGAGATAAATGAAAAAGTTTCACAGAGTTACTTTTGGCAATTACCAATTGTGAGGAG 60
 QY 108 AtgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGly 127
 Db 61 AGGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGTAAAGCCCTCCAAATAGACGTGGGC 120
 QY 128 CysProValLysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPhe 147
 Db 121 TGTCTCTGTGAAGTACAGCTGAGATCTGGGAGAGAAAAATTTCTTGAGTTGTACGCTTC 180
 QY 148 ArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeu 167
 Db 181 AGAGGACCCCTGTTAGCAGAGAGGACAGTCTCCGGAATATTTCTTGAGTTGAAATGTCTG 240
 QY 168 GluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGln 187
 Db 241 GAAGAAGGTGCTGTCAAGGTTTCACTGACGGGTGTACCAAGGGAACACAGCTTTTTCAG 300
 QY 188 CysAspGluAspCysGly----PheValAlaLeuAspLysLeuGluLeuIleGluAspAsp 206
 Db 301 TGTGATGAAGATTGTGGCGTGTGTTGTCATTGACCAAGCTAGCAACTCATAGAAGATGAT 360
 QY 207 AspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuPro 226
 Db 361 GACACTGCAATGGAAAGTATTACGAGGTCCTGGGACACACAAATGACAGTGCAGACTTCT 420
 QY 227 ProLeuGluLeuAsnSerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThr 246
 Db 421 CCTTTGGAAATAAATCCAGAGTTTCTTTGAAGGTTGGAGAAACAATAGATATCTGGAACA 480

```
QY 247 ValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAsp 266
Db 481 GTTATATTCTGTGATGTTTGCAGAAAGAAAGCTTAGGATATTGTTGTGTGGAC 540
QY 267 MetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPhe 285
Db 541 ATGGATAACCCATTATGGCAACTGGGATGGAAGATTGATGGAGCTTGAGCTTTT 600
QY 286 AlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThr 305
Db 601 GCCTGTGTTCAAAAGTAGTCAATCTTATTGACATCAATGATGATATCCAGAGAGTGTGACG 660
QY 306 GlnGluArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySer 325
Db 661 CAGGAAAGAGGCGCTCCCAAACTTGCCTTTATGTCAGAGGCTTGGGGACAAAGGTTCA 720
QY 326 SerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg---Arg 344
Db 721 TCCAGTCATAATNAACCAAGGCTACAGGATCTACCTCAGACCTCGAATAGAAACAGA 780
QY 345 SerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSer 364
Db 781 TCTGAATTATTTTATACCTTAAATGGTCTTCTGTGACTCACACCAACCAATCCAATCA 840
QY 365 LysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIle 384
Db 841 AAAAATACATGTTACATGATGAAGTTGCAGAAAGCCCTGCAAAATCTCTTACAGAGATA 900
QY 385 SerThrAspPheAspArgSerSerProLeuGlnProProValAsnSerLeuThr 404
Db 901 TCTACAGACTTTGACCGCTCTTCCACCACACTCCAGCCTCCTCTGTGAACCTCACTGACC 960
QY 405 ThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGly 424
Db 961 ACCGAGACAGATTCACACTCTTACCACTCAGTCTCACCAGAGTGCCCAATACCAATGGA 1020
QY 425 SerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThr 444
Db 1021 AGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAGTCTGTATGGAAGAGCTAAACACT 1080
QY 445 AlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGlu 464
Db 1081 GCACCCGTCACAGAGAGTCCACCTTTGGCCATCCCTCCTCGGAACTCACATGGTCTAGAA 1140
QY 465 ValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTyrPile 484
Db 1141 GTGGGCTCATTTGGCTGAGTTAGGAGAACCTTCTTCTATGGGGTAATCCGTTGGATC 1200
QY 485 GlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAla 504
Db 1201 GGTCCAGCCACCAGGACTGATGAAGTCTCGCTGGACTGGAACCTGGAAGATGAGTGTGCA 1260
QY 505 GlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAla 524
Db 1261 GGCTGTACGGATGGAACCTTTCAGAGGCACCTCGGTATTTCACCTGTGCGCCTGGAAGAGCG 1320
QY 525 LeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProVal 544
Db 1321 CTGTTTGTGAAACTGAAGAGCTGAGGCCCTGACTCTAGGTTTGCATCATTTGCAGCCGGTT 1380
QY 545 SerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValVal 564
Db 1381 TCCAAATCAGATTGAGCGCTGTAATCTTTAGCATTTTGGAGGCTACTTAAGTGAAGTAGTA 1440
QY 565 GluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLys 584
Db 1441 GAAGAAAATACTCCACCAAAATGGAAGAAAGAGGCTTGAGAGATAATGATTGGGAAGAG 1500
QY 585 LysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPhe 604
Db 1501 AAAGGCATCCAGGCTCATTAACAATCTTGTATTCTTAGACTCAACCTTATTCTGCTTATT 1560
QY 605 AlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGlu 624
```

```
Db 1561 GCITTTAGTTCTGTTCTGGACACTGTGTACTTTAGACCCCAAGAAAGACGATGTAGAA 1620
QY 625 TyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyr 644
Db 1621 TATTATAGTGAAACCCCAAGAGCTTACTGAGCAGACAGAAATTTGTTAATCCTCTGAGAATATAT 1680
QY 645 GlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAla 664
Db 1681 GGATATGTGTGTGCCCAAAATTTATGAAACTGAGGAAATATCTTGAAGAGGTGGAGGCT 1740
QY 665 AlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIle 681
Db 1741 GCATCAGGATTTCCTCTGAAGAAAAGGT-----GACCATCTTAACTTA 1785
```

Search completed: April 18, 2005, 15:55:04
Job time : 1160 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:26:40 ; Search time 8972 Seconds

(without alignments)
5125.279 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSFYWEERI.....RLLCDAYMCYQSPWMSLYK 949

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/uspto.spool.p/US09671687/runat.18042005.115018.17125/app.query.fasta_1.1095
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09671687 @CGN 1.1.5932 @runat.18042005.115018.17125 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4983	99.0	3302	6	CQ834224 Sequence
2	4983	99.0	3540	9	BC012342 Homo sapi
3	4983	99.0	5414	9	AB020656 Homo sapi
4	4980	98.9	3302	6	CQ834222 Sequence

5	4971.5	98.8	5371	9	HS250014
6	4968.5	98.7	3311	6	CQ834226
7	4962	98.6	3480	6	CQ719792
8	4942	98.2	4527	6	BD231207
9	4718.5	93.7	4314	10	AK122389
10	4718.5	93.7	4501	10	BC042438
11	4712	93.6	3267	10	BC082001
12	4601	91.4	2845	6	BD160617
13	4601	91.4	2845	6	AX883937
14	4601	91.4	2845	9	AK024348
15	4600.5	91.4	241990	2	AC123449
16	4600.5	91.4	251132	2	AC098162
17	3676	73.0	2523	6	AR338799
18	3433.5	68.2	2341	9	AK000187
19	3022.5	60.0	2569	6	BD160470
20	3022.5	60.0	2569	6	AX883681
21	3022.5	60.0	2569	9	AK024212
22	2930	58.2	1954	9	AK056226
23	2493	49.5	241630	2	AC131881
24	2468	49.0	2116	6	BD135433
25	2468	49.0	2116	6	CQ867781
26	2468	49.0	2116	6	AR534837
27	2468	49.0	2116	6	AX017216
28	2108.5	41.9	182059	2	AC133651
29	1736	34.5	251957	2	AC126867
30	1662	33.0	2426	10	BC049879
31	1442.5	28.7	212531	2	AC145178
32	1430.5	28.4	129025	2	AC145180
33	1370	27.2	1069	5	BX933038
34	1344	26.7	1335	5	BX934133
35	1262	25.1	837	6	BD149741
36	1262	25.1	837	6	AX869679
37	1215	24.1	758	6	BD147076
38	1215	24.1	758	6	AX867014
39	1132.5	22.5	2068	3	AY071592
40	1111	22.1	151567	2	AC145503
41	1111	22.1	199277	2	AC145445
42	1109.5	22.0	208061	2	AC145254
43	1108.5	22.0	168271	9	AC007728
44	1097.5	21.8	141663	2	AC145018
45	1095.5	21.8	163319	9	HS303140

ALIGNMENTS

RESULT 1	CQ834224	Sequence 95 from Patent WO2004058805.	3302 bp	DNA	linear	PAT 29-JUL-2004
LOCUS	CQ834224					
DEFINITION	CQ834224					
ACCESSION	CQ834224					
VERSION	CQ834224.1	GI:50833761				
KEYWORDS						
SOURCE						
ORGANISM		Homo sapiens (human)				
REFERENCE						
AUTHORS		Matsuda A. and Yoneta S.				
TITLE		T cell activating gene				
JOURNAL		Patent: WO 2004058805-A 95 15-JUL-2004; - late				
FEATURES						
source		Asahi Kasei Pharma Corporation (JP)				
		Location/Qualifiers				
		1..3302				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				
		243..3104				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="CAH05327.1"				
		/db_xref="GI:50833762"				
		/translation="MSSGLWSQEKVTSFYWEERIFVLLLOQCSVTDKQTKLLKVPKG				
		SIQGYIQDRSVGHSRIPSAKGNKKNIGLKLEQPHAVLFVDEKDVNEINEKFTLELLA				

ITNCRFSLFKNRNLSKGLQIDVGCPUKVLQSRSEKPGVVRPRGPIAERTVSG
 IFPGVLLBEGRGQGTGVYQKQLFCQDEGCVFVALDKLELIEDDDTALESVDAG
 PFGWQVLEPPLEINSRVKSGVETIESGTVFCDVLPKGSILGYVGDPMNPIGNW
 DGRFDVQVCSFACVSTILLIINDIIPESVTQERRPPKLPFMRSGVDGKSSSHNKP
 KATGSDPGRNRSSELFITLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISDF
 DRSSPLQPPVNSLTTERPHSLPLSLTKMNTNGSIGHSPISLSAQSVMEELNAP
 VQESPPLAPGNSHGLEVSLAEVKNENPFYVIRWIQPPGLNEVLAGELEDESCA
 GCTDGRTRVFTCALKALFVKLSKCRPSRFSASLQVSNQIERCNSLAFGGYLS
 VVENETPPMEKEGLEIMIGKKGIQGHYNSCYLDSTLFLFAFSSVLDVILRPXK
 NOVEYTSQTELLRLTEINPLRIYGVYCATKIMKRLKLEKVEAAGSFTSEKDPPEF
 LNLPHILRLVPELRIRAGAKQVOCYFVQIFMEKNEKVGVPTIOQLLEWFSINSL
 KEAEPSCILIIQMPRGKDPKPKIIPSLLEINITLLEDTPROCRICGLWAYECRE
 CDDPDISAKIKQFCNTQVHLHPKRLNHNKINPVLKDLDPDWDWHGHCIPQCNM
 ELFALVCITSHYVAFVKYKDDSAWLFEDSMADRDGGQGNFIPOVTPCPVEGYLK
 MSLEDLHSLDSRRIOGCARRLLCDAWCMYQSPMTSLYK"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3302
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 6 Gaps: 4

US-09-671-687A-3 (1-949) x CQ834224 (1-3302)

QY 1 MetSerSerGlyLeuThrPserGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
 |||||
 Db 243 ATGAGTTACGGCTTATGGAGCCAAAGAAAGTCACTTCACTTACCTTGGGAAGCGGATT 302
 |||||
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 |||||
 Db 303 TTTTACTTCTTCAAGATCCAGCGTTACAGACAAACACACAAAGCTCCTTAAA 362
 |||||
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 |||||
 Db 363 GTACCGAAGGAGTAGTACAGATGATATTTCAAGATCGTTCTGTGGCGCATTTCAAGGATT 422
 |||||
 QY 61 ProSerAlaLysGlyLysLysAenGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 |||||
 Db 423 CCTTTCGAAAAGGCAAGAAATACAGATTGGATTAAATTTAGAGCAACCTCATGCA 482
 |||||
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 |||||
 Db 483 GTTCTCTTTGTGATGAAAGGATGTTGTAGATATAATGAAAGTTTACAGAGTTACTT 542
 |||||
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 |||||
 Db 543 TTGGCAATTACCAATTTGTGAGGAGAGGTTCCAGCTGTGTTAAAAACAGAACAGACTAAGT 602
 |||||
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 |||||
 Db 603 AAAGGCTTCAATATAGACGTGGCGCTCTCTGTGAAAGTAGTACAGTGTGAGATCTGGGGAGAA 662
 |||||
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 |||||
 Db 663 AAATTTCTGGAGTTGTACGCTTACAGAGNCCCTGTTAGCAGAGAGGACAGTCTCCGGA 722
 |||||
 QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179
 |||||
 Db 723 ATATTCTTTGGAGTTGAAATGCTTGGGAAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 782
 |||||
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 |||||
 Db 783 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTGATTTGGAC 842
 |||||
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 |||||
 Db 843 AAGCTAGAACTTAGAAGATGATGACCTGCAATTGGAAAGTATTACCGAGGCTCTGGG 902
 |||||
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 |||||

Db 903 GACACAATCGAGGTCGAACCTTCTCTCTTGGAAATAAACTCCAGAGTTCTTTTGAAGGTT 962
 |||||
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 |||||
 Db 963 GGAGNAACAATAGAACTCGAACAGTTATATCTGTGATGTTTTCGACAGAAAGAAGC 1022
 |||||
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 |||||
 Db 1023 TTAGGATATTTTGTGTGGTGTGGACATGGAATAACCTATTGGCAACTGGGATGGAAGATT 1082
 |||||
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 |||||
 Db 1083 GATGGAGTGCAGCTTTGTAGTTTGTGCTGTGTTGAAAGTACAATTTTATGGCATCAAT 1142
 |||||
 QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
 |||||
 Db 1143 GATATCATCCAGAGAGTGTGACGAGGAAGAGAGGCTCCCAAACTTGTGCTTTATGTCA 1202
 |||||
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 |||||
 Db 1203 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 1262
 |||||
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 |||||
 Db 1263 TCAGACCTCGAAATAGAAACAGATCTGAATTTTATACCTTAAATGGGTCTCTGTT 1322
 |||||
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp 376
 |||||
 Db 1323 GACTCAACACCAACATCCAAATCAAAAATACATGTTACATTTGATGAGTTGCAGAAGC 1382
 |||||
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
 |||||
 Db 1383 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCTCCAG 1442
 |||||
 QY 397 ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerIleu 416
 |||||
 Db 1443 CCTCTCTCTGTGAACCTCACTGACCACCGAGAACAGATTCACACTCTTTTACCATTTCAGTCTC 1502
 |||||
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 |||||
 Db 1503 ACCAATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTACGCCAG 1562
 |||||
 QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 |||||
 Db 1563 TCTGTAAATGGAGAGCTAAACACTGCGCCGTCACAGAGAGTCCACCTTGGCCATGCT 1622
 |||||
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 |||||
 Db 1623 CTTGGGAACTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCT 1682
 |||||
 QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 |||||
 Db 1683 TTTCTATGGGGTAAATCCGTTGGATCGGTGACCGACCCAGGACTGAATGAAGTGTCTCGCTGGA 1742
 |||||
 QY 497 LeuGluLeuLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 |||||
 Db 1743 CTGGAATCGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTAT 1802
 |||||
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
 |||||
 Db 1803 TTCACCTGTGCTCGCCCTGAAGAAGGCGCTGTTTGTGAAACTGAAGAGCTGACGCGCTGACTCT 1862
 |||||
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 |||||
 Db 1863 AGGTTTGGCATCATTTGACGCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1922
 |||||
 QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
 |||||
 Db 1923 GAGGAGTACTTAAGTAGAGTAGTAGAAGAAAATACTCCACAAAATAATGGAAAAGAAGGC 1982
 |||||
 QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
 |||||
 Db 1983 TTGGAGATAATGATTGGGAAGAAGAAGGCATCCAGGTCATTACAATTTCTTGTACTTATA 2042
 |||||

QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
 DB 2043 GACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGACACTGTGTTACTTTAGA 2102
 QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
 DB 2103 CCCAAGAAGAACGATGATATATATAGTGAACCCAGAGCTACTGAGACAGAA 2162
 QY 637 lleValAsnProLeuArgileTyrGlyTyrValCysAlaThrLysileMetLysLeuArg 656
 DB 2163 ATTGTTAATCTCTCAGAAATATATGATATGTTGTGCCACAAAATTTATGAACCTGAG 2222
 QY 657 LysileLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu 676
 DB 2223 AAAATCTTGAAGAGTGGAGCTGCATCAGATTTTACCTCTGAAGAAAAGATCTCTGAG 2282
 QY 677 GluPheLeuAsnileLeuPheHisileLeuArgValGluProLeuLeuLysileArg 696
 DB 2283 GAATTCCTGAATATCTGTTTCATCATATTTTAAGGTAGAACCTTTGCTTAAAAATAGA 2342
 QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
 DB 2343 TCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAATTTTATGGAATAATGAG 2402
 QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrPheIleAsnSerAsnLeu 736
 DB 2403 AAAGTTGGCGTTCCCAACATTCAGCAGTTGTTAGATGCTCTTTTATCAACAGTAACTG 2462
 QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
 DB 2463 AAATTTGAGAGGACCAATCATGCTGTGATTTTCAGATCCCTCGATTTGGAAAAGACTTT 2522
 QY 757 LysLeuPheLysLysilePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
 DB 2523 AAATTTTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTTACTTTGAAGAC 2582
 QY 777 ThrProArgGlnCysArgileCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
 DB 2583 ACTCCACAGACGTCCGATATGTGAGGCGTTCGAATGTATGAGTGTAGAGAACTCTAC 2642
 QY 797 AspAspProAspIleSerAlaGlyLysileLysGlnPheCysLysThrCysAsnThrGln 816
 DB 2643 GACGATCCGGACATCTCAGCTGGAATAATCAACGAGTTTTGTAAAAACCTGCAACACTCAA 2702
 QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
 DB 2703 GTCCACCTTCATCCGAGAGGCTGAATCATATAATATAACCCAGTGCATCTCCCAAGAC 2762
 QY 837 LeuProAspTrpAspTrpArgHisGlyCysileProCysGlnAsnMetGluLeuPheAla 856
 DB 2763 TTACCCGACTGGAGTGGAGACACGCTGCATCCCTTGCAGAAATATGAGTTATTTGCT 2822
 QY 857 ValLeuCysileGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 DB 2823 GTTCTCTGATAGAACCAAGCCACTATGTGTTTGTGAAGTATGGGAGACGATTTCT 2882
 QY 877 AlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
 DB 2883 GCCTGGCTCTCTTTGACAGCATGGCCGATGGGATGGTGGTGCAGAAATGGCTTCAACATT 2942
 QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 DB 2943 CCTCAAGTCAACCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGAGACCTTG 3002
 QY 917 HisSerLeuAspSerArgArgileGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 DB 3003 CATTCCTTGGACTCCAGGAGATCCAAAGCTGTGCACGAGACTGCTTGTGTGATGCATAT 3062
 QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 DB 3063 ATGTGCATGTACCAGATCCAAACATGAGTTGTGTACAAA 3101

RESULT 2
 BC012342
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3540)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zerbberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Uesdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.,
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) - late
 22389257
 12477932
 2 (bases 1 to 3540)
 Strausberg,R.
 Direct Submission
 Submitted (15-AUG-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-x@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjelli, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalush, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 29 Row: c Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 1416257.
 Location/Qualifiers
 1. .3540
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:19923 IMAGE:4552767"
 /tissue_type="Uterus, leiomyosarcoma"

FEATURES
 source

gene	/clone_lib="NIH_MGC_46"		Db			517	GTTCTCTCTTGTGTGATAAAAGGATGTTGTGTAGAGATAAATGAAAAAGTTTACACAGAGTTACTT		576
	/lab_host="DH10B-R"								
gene	/note="Vector: pOTB7"		Qy			100	LeuAlaIleThrIleThrCysGluGluUaArgPheSerLeuPheIysAsnArgAsnArgLeuSer		119
	1. .3540								
CDS	/gene="CYLD"		Db			577	TTGCAATATACCAATTTGTGAGGAGAGGTTGAGCTGTTTAAAAACAGAAAACAGACTAAGT		636
	/note="synonyms: CDWT, CYLD1, EAC, KIAA0849, HSPC057"								
CDS	/db_xref="LocusID:1540"		Qy			120	LysGlyLeuGlnIleLeuValGlyCysProValIysValGlnLeuLeuArgSerGlyGluGlu		139
	/db_xref="MIM:605018"								
CDS	277. .3138		Db			637	AAAGCCCTCCAAATAGACGTGGCTGCTGTGAAGTACAGCTGAGATCTGGGGAGAA		696
	/gene="CYLD"								
CDS	/codon_start=1		Qy			140	LysPheProGlyValValAlaArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly		159
	/product="CYLD protein"								
CDS	/protein_id="AAH12342.1"		Db			697	AAATTTCTCGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGTCTCCGGA		756
	/db_xref="GI:15214434"								
CDS	/db_xref="LocusID:1540"		Qy			160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal		179
	/db_xref="MIM:605018"								
CDS	/translation="MSGSLWSQEKVTPYWEERIFYLLQBCSVTDKQTKLVKPKG SIGQIQRSVGHRIKISAKGKQIQLKILEQPHAVLFVDEKDVVEINEKFTLELLA ITNEERPSLPKRNRLSKGQIDVGCQPKVQLRSBEKFPVVRFRGPLLAERTVSG IFFGVELLEGRGQFTDVGQKQLFCQDEDCGVFVALDKLELIEDDTALESVDYAG PGDMQVPLEINSLRVSLKGTIBSGVTIFCDVLPKGESLGFVGVDMDPIGNW DGRDVGQVLCSPACVESTILHINDIIPESVTQERRPPKLAFMGRGVGDKGSSSHNP KATGSTDGPNRNSRLPYTLNGSSVDSQPSKSNWTYIDEVAEDPAKSLTEISTDP DRSPGPPQPPVNSLTITENRPHSLPFLSLTKMPNTNGSIGHSPLSAQSVMEELNAP VQSPPLAMPNGSHGLEVSLAENPNPPYGIWRIGQPPGLENLVAGLELEDECA GCTDGTGRTRYFTCAIKKALFVKLSCRPSRPSASLPQVSNQIERCNSLAFGGYLS VVENTPPKMEKLEIMIGKKKGIQGHYNSCYLDSTLFCFLFAFSSVLDTVLLRPKEK NDVYSSETOLLERILNPLRIYGVVCATKIMKLRIKLEKVEAASGFTSEKDPSEF NLILFPHILVEPLKIRSAQKVQDCVFYQIFMEKNEKVGVTIQQLLEWSFINSNL KFAPASCLLIIQPRFGQDFKFKIIPSLNELNTDILEDTPRQICGGLWAMEYCRE CYDDPISAGIKIQKFCNTQVHLKRLNHNKYNPSLPKDLDPDWDRHGICPCQNM ELFALVCIETSHYVAFVKYKQDQSAWLPFDQSMADRDGQNGFNIPQVTPCPVEGYLK MSLEDHSLDSRRITQGCARRLLCDAYMCYQSPTMSLYK"		Db			219	AspThrMetGlnValGluLeuProProLeuIleAsnSerArgValSerLeuIysGly		238
	745. .867								
misc_feature	/gene="CYLD"		Qy			259	LeuGlyTyrPheValGlyValAlaPheMetAspAsnProIleGlyAsnTrpAspGlyArgPhe		278
	/note="CAP_GLY; Region: CAP-Gly domain. CAP stands for cytoskeleton-associated proteins"								
misc_feature	/db_xref="CDD:pfam01302"		Db			1057	TTAGGATATTTTGTGTGTGGACATGATGATAACCTATTGGCACTGGGATGGAAGATT		1116
	1720. .2025								
misc_feature	/gene="CYLD"		Qy						

Db 1597 TCTGTAATGGAAGAGCTAAACACTGTCACCCGCTCCRAAGAGAGTCCACCCCTTGGCCATGCCT 1656

QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476

Db 1657 CTTGGGAATCATAGTGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCCT 1716

QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496

Db 1717 TTCTATGGGGTAATCCGTTGGATCGGTGAGCCACACAGGACTGAATGAAGTGTCTCGCTGGA 1776

QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516

Db 1777 CTGGAACCTGGAAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGGCACTCGGTAT 1836

QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536

Db 1837 TTCACCTGTGCGCTGAAGAAGCGCTGTTTGTGAACCTGAAGAGCTGCAGGCGCTGACTCT 1896

QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556

Db 1897 AGGTTTGCAATCATGAGCCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1956

QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576

Db 1957 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAGAGGC 2016

QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596

Db 2017 TTGGAGATAATGATTGGGAAGAAGAACATCCAGGGTCATTACAAATCTTGTACTTAA 2076

QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616

Db 2077 GACTCAACCTATTCTGCTATTGCTTTTGTAGTTCTGTTCGACACTGTGTACTTAGA 2136

QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636

Db 2137 CCCAAGAAAAGAACGATAGAAATATATAGTGAACCAAGAGCTACTGAGGACAGAA 2196

QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656

Db 2197 ATTGTTAATCCTCTGAGATATATGATATGCTGTGCCACAAAATTTATGAACCTGAGG 2256

QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu 676

Db 2257 AAAATACTTGAAGAGTGGAGGCTGATCAGGATTTACCTCTGAAGAAAAGATCCTGAG 2316

QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696

Db 2317 GAATTTCTTGAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGA 2376

QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716

Db 2377 TCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGAAAAAATGAG 2436

QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736

Db 2437 AAAGTTGGCGTCCCAACATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACCTG 2496

QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756

Db 2497 AAAATTTGAGAGGACCATCATGCTCTGATTAATTCAGATGCTCGATTTGGAAAAAGACTTT 2556

QY 757 LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAsp 776

Db 2557 AAACATTTTAAAAAAATTTTCTCTCTGGAATTAATAATACAGATTTACTTTGAAGAC 2616

QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796

Db 2617 ACTCCAGACAGTCCGGATATGTGAGGCGTTCGATGTATGAGTGTAGAGATGCTAC 2676

QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816

Db 2677 GACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2736

QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836

Db 2737 GTCACCTTTATCCGAAGAGGCTGAATCATAAATATATACCCAGTGTCACTTCCCAAGAC 2796

QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856

Db 2797 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTGCT 2856

QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876

Db 2857 GTTCTCTGCATAGAAAACAGCCACTATGTTGCTTTTGTGAAGTATGGAGAGCAGATTCT 2916

QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896

Db 2917 GCCTGCTCTCTTTTGCACGATGGCCGATCGGATGGTGGTCAAGATGGCTTCAACATT 2976

QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916

Db 2977 CCTCAAGTACCCCATCCAGAGTAGAGAGTACTTTGAAGATGTCTCTGGAAGACCTG 3036

QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936

Db 3037 CATTCCTTGGACTCCAGGAGATCCCAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3096

QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949

Db 3097 ATGTGCATGTACCAGAGTCCCAACATGAGTTGTGTACAAA 3135

RESULT 3

AB020656 5414 bp mRNA linear PRI 17-MAY-2001
LOCUS Homo sapiens mRNA for KIAA0849 protein, partial cds.

DEFINITION AB020656

ACCESSION AB020656

VERSION AB020656.2 GI:14133220

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases)

AUTHORS Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirose,M., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.

TITLE Prediction of the coding sequences of unidentified human genes.

XII. The complete sequences of 100 new cDNA clones from brain which

code for large proteins in vitro

JOURNAL DNA Res. 5 (6), 355-364 (1998)

MEDLINE 99156230

PUBMED 10048485

REFERENCE 2 (bases 1 to 5414)

AUTHORS Ohara,O., Suyama,M., Kikuno,R., Nagase,T. and Ishikawa,K.

TITLE Direct Submission

JOURNAL Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute,

Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba

292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,

Fax:+81-438-52-3914)

COMMENT On May 17, 2001 this sequence version replaced gi:4240186.

FEATURES

source

1. .5414

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="fh04363"

/sex="male"

/tissue_type="brain"

/clone_lib="pBluescriptII SK plus"

/dev_stage="adult"

/notes="vector:pBluescriptII SK plus. This sequence is

replaced that of hk05904 cDNA as a representative cDNA

sequence for KIAA0849"

1. .5414

/genes="KIAA0849"

<426. .3308

gene

CDS

NOT LATE

/gene="K1AA0849"
 /codon_start=1
 /product="K1AA0849 protein"
 /protein_id="BAA74872.2"
 /db_xref="GI:14131221"
 /translation="VFVNINIMSSGLNSQEKVTSYWEERIFYILLQBCSVTDKQTK
 LKVPKSGISQYIQRSVGHSPKSAKKNQIGLILQPHAVLFLVDEKQVVEINEK
 FTELLALATNCEERFLSKNRNLSGLQIDVQKVPQVLRSGEEKFPQVFRPGLL
 AERTVSGIFGVVLEERQGGQFTDGVYQKQFQDCEQDFVAFDLKLELDDDTA
 LESDVAGQDTPVLEINRSVLKVGETIESGTVIFCDVLPKESLGVFVGDM
 DNPIGNMDRPGDVLCPACVSEFILLHINDIIPESVTOERPPKLAIFMSRGVGDGK
 SSSHKNKATGTSIPDGNRNSLEFYTLNGLSSVDSPQSKSNKNTYIDEVAEDPAKSL
 TEISDFRSPSSLPQPPNLSLITNRFHSLPFLTKMPTNGTSGHSPLSLSAQSVM
 BELNTAPQESPLAMPNGSHGLEVAEKENPFYVIRWIGQPGLENEVLGL
 ELEDGCACTDCTFGTRYFTCALKKALFVKLSCRPSRPFASLPQVNOIERCNSLA
 FGGYLSVEVENTPKMEKEGLEIMGKKKGIOGHVNSCYLDTSLFCFLPFRSSVLDTV
 LLRPKENDVEYSETQELLREIYNPLRIYGVCAKIMKRLKILEKVEAASGTISE
 EKDPBEFLNHHILRLKISAGQKQVQDIFYQIFMEKNKVKVPTTQQLEW
 SFINSNLPFAEAPSLIITQMPFRGDFKLFKKIFPSLELNTDLDLTPRCQRI CGGL
 AMYECRECVDDPDISAGKIKQPKCTCNTQHLHPKRLNKHYPVSLPKDLDPDWRHG
 CIPCONMELFVLCITETSHYVAFVKYKDDSAWLPFDSMADRDGGQNGNIPQVTPCP
 EVGEYIKWSLEDLHSLDSRRIQGCARRLLCDAYMCMYQSPMTSLYK"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 5414
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 9 Gaps: 4

US-09-671-687A-3 (1-949) x AB020656 (1-5414)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTrpTrpGluGluArgIle 20
 DB 447 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCACTCCCTTCTGGGAAGACGGATT 506
 QY 21 PheTyLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 507 TTTTACTTCTCTTCAAGAATGACGCGTTACAGACAAAACAAACAAAGCTCTCTTAA 566
 QY 41 ValProLysGlySerIleGlyGlnTyrlleGlnAspArgSerValGlyHisSerArgIle 60
 DB 567 GTACCGAAGGAAGTATAGACAGTATATTCAAGATCGTTCCTGTGGGCATTCAGGATT 626
 QY 61 ProSerAlaLysGlyLysLysAenGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 DB 627 CCTTCTGCAAAAGGCAAGAAATACAGATTGGATTAAAAATTTAGAGCAACCTCATGCA 686
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
 DB 687 GTTCTCTTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGATTACTT 746
 QY 100 LeuAlaIleThrAsnCySGLuGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 747 TTGGCAATTACCAATTGTGGAGAGAGTTGAGCTGCTGTTTAAAAACAGAAACAGACTAAGT 806
 QY 120 LysGlyLeuGlnIleAspValGlyCySProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 807 AAAGGCTTCAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 866
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 867 AATATTTCTGGAGTTGTACGCTTTCAGAGGACCCCTGTTAGCAGAGAGGACAGCTCCGGA 926
 QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
 DB 927 ATATTCTTTGGAGTTGAAATGCTGGAGAAGAGTTCGTGGTCAAGTTTCACTGACGGGGTG 986
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCySAspGluAspCySGLy---PheValAlaLeuAsp 198
 DB 987 TACCAAGGGGAAACAGACTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTTGTGATTCGAC 1046

QY 199 LysLeuGluLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 1047 AAGCTAGAACTCATAGAAGATGATGACATGCACTGCAATGGAAGTATACGAGGCTCTCTGG 1106
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 1107 GACACAATGCAAGTGAACCTTCTCTCTTGGAAATAAATCCAGAGTTCTTTTGAAGGTT 1166
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 1167 GGAGAAACAATAGAACTCGAACAGTTATATCTGTGATGTTTGGCAGGAAAAGAAAGC 1226
 QY 259 LeuGlyTyPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1227 TTAGGATAATTTTGTGTGGTGGACATGGATAACCTTATTGGCAACTGGGATGAAGATT 1286
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1287 GATGGAGTCAGCTTTGTAGTTTGTGCTGTGTGAAAGTACAATTTCTATTGGCATCAAT 1346
 QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
 DB 1347 GATATCATCCAGAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1406
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 1407 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC 1466
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyThrLeuAsnGlySerSerVal 356
 DB 1467 TCAGACCTCGGAAATAGAAACAGACTCTGAATTTATATACCTTAAATGGGTCTCTCTGT 1526
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrlleAspGluValAlaGluAsp 376
 DB 1527 GACTCAACACCAATCCAAATCAAAATAATACATGATGATGATGATGATGATGATGATGAT 1586
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
 DB 1587 CTTGCAAAATCTCTACAGAGATATCTACAGACTTTGACCGTCTCTTCAACCACTCCAG 1646
 QY 397 ProProProValAsnSerLeuThrGluLysArgPheHisSerLeuProPheSerLeu 416
 DB 1647 CTTCTCTCTGTGAACCTCACTGACACCGAGAGACAGATTCACACTCTTTACCACTTCTC 1706
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 1707 ACCAAGATGCCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTCTCTGTGAGCCAG 1766
 QY 437 SerValMetGluGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 1767 TCTGTATGGAAGAGCTTAAACACTGACCCGTCAGAGAGTCCACCTTGGCCATGCCT 1826
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 1827 CTTGGGAACCTCACATGCTCTAGAAGTGGGCTCATTTGGCTGAAAGTAAAGGAAACCTCT 1886
 QY 477 PheTyGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 1887 TTTCTATGGGGTAATCCGTTGGATCGGTTCAGCCACAGGACTGAAATGAAGTCTCGCTGGA 1946
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516
 DB 1947 CTGGAACCTGGAAGATGATGTCAGGCTGTACGGATGGAACCTTCAGAGGACCTCGGTAT 2006
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
 DB 2007 TTTCACTGTCCCTGGAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCTGACTCT 2066
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 DB 2067 AGGTTTTCATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTGTAACCTCTTTAGCATTT 2126

```
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 2127 GGAGGCTACTTAAGTAGAGTAGAGAAATACTCCACCAAAATGGAAAAGAGGC 2186
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 2187 TTGGAGATAATGATTGGGAAGAAGAGGATCCAGGGTCATTACAAATCTCTGTACTTA 2246
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2247 GACTCAACCTTATCTGCTTATTGCTTTAGTTCTGTTCTGACACTGTGTACTTAGA 2306
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuArgThrGlu 636
Db 2307 CCCAAAGAAAGACGATGAGATATTATAGTGAACCCCAAGAGCTACTGAGACAGAA 2366
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2367 ATTGTTAATCTCTGAGATATATGATATGTTGTGCCACAAAATTTATGAACTGAGG 2426
QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu 676
Db 2427 AAAATACCTTGAAGAGGTGGAGCTGCATCAGGATTTACCTCTGAGAAAGATCCTGAG 2486
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2487 GAAATCTTGAATATCTGTTTCATCATATTTTAAGGGTAGACCTTTGCTAAAAATAAGA 2546
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2547 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAATTTTATGAAAAAATGAG 2606
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2607 AAAAGTTGGCGTTCCCACAATTCAGCAGTTGTTAGAAATGCTTTTATCAACAGTAACCTG 2666
QY 737 LysPheAlaGluAlaProSerCysIleuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2667 AAATTTGAGAGGACCAATCATGCTCTGATATTTCAGATCCCTCGATTTGGAAGACCTTT 2726
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuAsp 776
Db 2727 AAACATTTTAAAAAATTTTCTCTCTGGAATTAATAATACAGATTACTTTGAAGAC 2786
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2787 ACTCCACAGACAGTCGGGATATGTCGAGGCTTGCAATGTATGAGTGTAGAGATGCTAC 2846
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2847 GACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAAAACCTGCAACACTCAA 2906
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2907 GTCCACCTTCATCCGAAGAGGCTGAATCAATAATAATAACCCAGTGTCATCTCCCAAGAC 2966
QY 837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2967 TTACCCGACTGGGACTGGAGACACGGCTGATCCCTTGCAGAAATATGAGTTATTGCT 3026
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
Db 3027 GTTCTCTCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGACGATTC 3086
QY 877 AlaTyrLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 3087 GCCTGGCTCTCTTGTGACAGCATGCGCATCGGATGGTGGTCAGAAATGGCTTCAACATT 3146
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 3147 CCTCAAGTACCCCATGCCAGAGTAGAGAGTACTTGAAGATGTTCTCTGGAAGACCTG 3206
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAlaTyr 936
```

```
Db 3207 CATTCTTGACTCCAGAGATCCAAAGCTGTGCACGAACTCTTTGTGATGATAT 3266
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3267 ATGTGCAATGTACCAGAGTCCAAACAATGAGTTGTACAAA 3305

RESULT 4
CQ834222
LOCUS CQ834222
DEFINITION Sequence 93 from Patent WO2004058805.
ACCESSION CQ834222
VERSION CQ834222.1 GI:50833759
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Matsuda, A. and Yoneta, S.
TITLE T cell activating gene
JOURNAL Patent: WO 2004058805-A 93 15-JUL-2004; -late
Asahi Kasei Pharma Corporation (JP)
FEATURES
source 1..3302
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
243..3104
/note="unnamed protein product"
/codon_start=1
/protein_id="CAH05326.1"
/db_xref="GI:50833760"
/translations="MSSGLWSQEKVTSPPYWEERIFYLLLOBQSVTDKOTOKLLKVPKG
SIGQYIQDSVGHSHRIPSAKKNOIGLKILEOPHAFVFDKDVVEINEKFTLLA
IINCEERFSLFNKNRSLKGLQIDGCPVKQLRSGEKFPVVRFGPLAERTVSG
IFPGVLLBEGRGQGTGVYQKQLFQDDEDCGVFALDKLELIEDDDTALSDYAG
PDGMQVLEPLPEINRSVLKGETIESGTIVICDVLFGKESLGYFVGVDMDNPTGNW
DGRFDVQLCSFACVESTILLHINDIIPESVTOERRPKLAFMRSGVDGKSGSHNKP
KATGSTDPGNRSLFYTLNGSSVSDSPQSKSNWTYIDEVAEDPAKSLTETSTDF
DRSSPLQPPVNSLTENRFHSLPFLTKMPTNGSIGHPSLSIQAQSVMEELNAP
VQESPPLAMPNGSHGLEVSLAEVKENPFYGVIRWIGQPPGNEVLAGELEDECA
GCTDGRFTGYFTCAKALFVKLSRPSDPSASIQVSNQIERNCSLAFGGYLSL
VVENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTFLCFLAFSSVLDTVLLRPEK
NDVEYSETLELRTIVNPLRYGVCAIKMLKLEKVEAASGFTSEKQDPEEF
LNLPHILRVEPLLIARSAGQVQDCYFQIEMEKNEKVGVPITIQQLLEWFSINL
KFAEAPSCILLIQMPREGKDKPLFKLIPPSLELNTDLLEDTPRCRCICGGLAMECRE
CYDDPDISAGIKQPCNTQVHLHPKRLNHNKYNPSLPKDLDFDMDWRHGCICQNM
ELFALVICIETHYVAFVKYKQDSAWLFFDSMADRDGGQNGFNIPQVTPCPVEGYLK
MSLEDLHLSRRRIQGCARRLLCDAYMCMYQSPMTSLYK"
ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3302
Score: 4980.00 Matches: 947
Percent Similarity: 99.48% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 98.93% Indels: 4
DB: 6 Gaps: 4

US-09-671-687A-3 (1-949) x CQ834222 (1-3302)
QY 1 MetSerSerGlyLeuTyrPsrSerGlnGluLysValThrSerProTyrTrrGluGluArgIle 20
Db 243 ATGAGTTTCAAGCTTATGGAGCCCAAGAAAAGTCACTTCACTCCCTACTGGGAAGACGGATT 302
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db 303 TTTTACTTGTCTTCTCAAGATTCAGAGTTCAGAGCAACAAACACAAACAAAGCTCTCTAAA 362
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 363 GTACCGAAGGAAGTAGTAGACAGTATATTCAAGATCGTTCTGTGGGCGCATTCAGGATT 422
```

QY	61	ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80	QY	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436
DB	423	CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAAAATTTAGAGCAACCTCATGCA	482	DB	1503	ACCAAGATCCCAATACCAATGGAGATTTGGCCACAGTCCACTTTCTGTGAGCCACG	1562
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99	QY	437	SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro	456
DB	483	GTTCCTCTTTGGTGAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTTT	542	DB	1563	TCGTGTAATGGAGAGCTAAACACTGACCCTGCTCAAGAGAGTCCACCTTTGGCCATGCCT	1622
QY	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119	QY	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476
DB	543	TTGGCAATTTACCAATTTGTGAGGAGAGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT	602	DB	1623	CCTGGGAACCTCAGATGGTCTAGAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCT	1682
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139	QY	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly	496
DB	603	AAAGGCTCCAAATAGACGTGGGCTGCTCTGTGAAAGTAGCTGAGATCTGGGGAGAA	662	DB	1683	TTCTATGGGGTAAATCCGTTGGATCGGTGAGCCAGCCAGAGACTGAATGAAGTGTCTCGCTGGA	1742
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159	QY	497	LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516
DB	663	AAATTTCTGGAGTTGTACGCTTCAGAGACCCCTGTAGCAGAGAGACAGTCTCCGGA	722	DB	1743	CTGGAACTGGAGAGATGAGTGTGAGGCTGTACGGATGGAAACCTTCAGAGGCACCTCGGTAT	1802
QY	160	IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal	179	QY	517	PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer	536
DB	723	ATATCTTTGGAGTTGAATGCTGGAGAGAGGTCTGGTCAAGTTCCTGACGGGGTG	782	DB	1803	TTCACTGTGCTGAGAGAGGCGCTGTTGTGAAACTGAGAGAGCTGCAGGCCCTGACTCT	1862
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198	QY	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556
DB	783	TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGCTATGGAC	842	DB	1863	AGGTTTGCATCATTTGCAGCGGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCATTT	1922
QY	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218	QY	557	GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly	576
DB	843	AAAGCTAGAACTCATAGAAGATGATGACACTGCATTGGAAAGTGATTACGCAGGCTCCTGGG	902	DB	1923	GGAGGCTACTTAAGTGAAGTAGAGAGAAATACTCCACCAATAATGGAAGAAAGAGGC	1982
QY	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238	QY	577	LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596
DB	903	GACCAATGCAGGTCGAACCTTCTCTTTGGAAATAAATCCAGAGTTCCTTTTGAAGGTT	962	DB	1983	TTGGAGATTAATGATTCGGAGAGAAAGAGCATCCAGGCTCATTTACAATTTCTTGTACTTA	2042
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258	QY	597	AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg	616
DB	963	GGAGAAACAATAGAACTGGAACAGATTATATCTGTGATGTTTTCAGGAGAAAGAGC	1022	DB	2043	GACTCAACCTTATTTCTGCTTATTTGCTTTTGTGTTCTGTTCTGGACACTGTGTTACTTAGA	2102
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278	QY	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
DB	1023	TTAGGATATTTTGTGTGGACATGATGATAACCTTATTTGGCACTGGGATGGAAGATT	1082	DB	2103	CCCAAGAAAGAAAGACGATGTAGAATATATATAGTGAACCCCAAGAGAGCTACTGAGACAGAA	2162
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297	QY	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656
DB	1083	GATGGAGTGCAGCTTTGTAGTTTTCGTGTGTTGAAAGTACAATTTATTTGCACATCAAT	1142	DB	2163	ATTGTTAACTCTCGAATATATGATGATGTGTGTGCCACAAAATAATATGAAGACTGAAG	2222
QY	298	AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer	317	QY	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
DB	1143	GATATCATCCAGAGAGTGTGACGACGAGAAAGAGGCGCTCCCAACTTGTGCTTTATGTCA	1202	DB	2223	AAAATACTTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAG	2282
QY	318	ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr	337	QY	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696
DB	1203	AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC	1262	DB	2283	GAATTTCTGCAATATCTGTTTCATCATATTTAAAGGGTAGAACCTTTGTGTAATAAAGA	2342
QY	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356	QY	697	SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu	716
DB	1263	TCAGACCCCTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGCTCTCTGTT	1322	DB	2343	TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATAG	2402
QY	357	AspSerGlnProGlnSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp	376	QY	717	LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu	736
DB	1323	GACTCAACCAACCAATCCAATCAAAAATACATGTTGATGATGAGTTGCAGAGAC	1382	DB	2403	AAAGTTGGCGTCCCACTTCCAGAGTGTGTAGATGGTCTTTTATCAACAGTAACCTG	2462
QY	377	ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln	396	QY	737	LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe	756
DB	1383	CCTGCAAAATCTCTACAGAGATATCTACAGACTTTTGACCGTCTTCACCCACCTCCAG	1442	DB	2463	AAATTTGCAGAGGCACCATCATGCTCATATTATTCAGATGCTCGATTGGAAAAAGACTTT	2522
QY	397	ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu	416	QY	757	LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAsp	776
DB	1443	CCTCCTCTGTGAACCTCACTGACCACCGAGAACAGATTCCACTCTTTTACCATTTCAGTCTC	1502	DB	2523	AAACTATTTAAAAAATTTTTTCTCTCGGAATTAATAATACAGATTTACTTTGAGAGAC	2582
				QY	777	ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr	796

Db	2583	ACTCCACAGACAGTCCCGGATATGTGGAGGCGCTTGCATATGAT
----	------	---

Db 872 ATATTTCTTTGGAGTTGAATTCCTGGAGAAAGTCTGGTGCACAGTTTCTACGTACGGGGT 931
QY TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 932 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTGATTTGGAC 991
QY LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 992 AAGCTAGAACTCATAGAAGATGATGACACTGCAATTGGAAAGTGATTAACGACGGTCTCGG 1051
QY AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 1052 GACCAATGCGAGTCCGAATTCCTCTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT 1111
QY GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 1112 GGAGAAACAATAGAACTGGACAGATTATATCTGTGATGTTTGGCAGGAAAGAAGC 1171
QY LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 1172 TTAGGATATTTTGTGGTGGACATGATGATTAACCTATTGGCAACTGGGATGGAAGATT 1231
QY AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1232 GATGGAGTCCAGCTTTGTAGTTTGGGTGTGTTGAAAGTACAATTCTATTGGCACAAT 1291
QY AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
Db 1292 GATATCATCCAGCTTTATCAGAGAGTGTGACGAGAAAGGAGGCTTCCAAACTTGCC 1351
QY PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr 334
Db 1352 TTTATGTCAAGAGTGTGGGGACAAAGGTTTCATCCAGTCATATAAACCAAGGCTACA 1411
QY GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
Db 1412 GGATCTACCTCAGACCTCGAAATAGAAACAGATCTGAATATTTTATACCTTTAAATGGG 1471
QY SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
Db 1472 TCTTCTGTGACTCAACACCAATCCAAATCAAAATAACATGGTACATTGATGAAGTT 1531
QY AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerPro 393
Db 1532 GCAGAGACCTCGAAATCTCTACAGAGATATCTACAGACTTGTACCGTCTCTCACCA 1591
QY ProLeuGlnProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
Db 1592 CCACTCCAGCTCCTCCTGTGAACCTCACTGACCACCGAGAACAGATTCACACTCTTTACCA 1651
QY PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
Db 1652 TTCAGTCTCCACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACACTTCTCTG 1711
QY SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeu 453
Db 1712 TCAGCCAGTCTGTATGGAAGAGCTTAAACACTGCACCCGTCCAGAGAGTCCACCCCTTG 1771
QY AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1772 GCCATGCCTCCTGGGAACCTCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAG 1831
QY AsnProProPheTyrGlyValIleArgTTPileGlyGlnProProGlyLeuAsnGluVal 493
Db 1832 AACCTCCTCTTCTATGGGGTAATCCGTTGGATCGGTGATCGGTGAGTGAATGAAGT 1891
QY LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db 1892 CTCCTGGACTGGAACCTGGAAGATGAGTGTGCAAGCTGTACGGATGGAACCTTCAGAGGC 1951
QY ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533

Db 1952 ACTCGGTATTTACCTGTGCCCCTGAAGAAGGCCCTGTTGTGAAACTGAAGAGCTGCAGG 2011
QY ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 2012 CCTGACTCTAGGTTTGCATCATTTGCAGCGGTTTTCCAAATCAGATTGAGCGCTGTAACCTCT 2071
QY LeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
Db 2072 TTAGCATTTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCANAATGGAA 2131
QY LysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSer 593
Db 2132 AAGAAGGCTTGGAGATAATGATTGGAGAGAAGAGGCATCCAGGCTCATTAACATCT 2191
QY CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2192 TGTACTTAGACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTCTGGACACTGTG 2251
QY LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu 633
Db 2252 TTACTTTAGACCCCAAGAAAAGACGATGTAGAAATATTATAGTGAACCCCAAGAGCTACTG 2311
QY ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
Db 2312 AGGACAGAAATTTGTAATCTCTCGAATATATGAGATATGTGTGTGCCACAAAATATTG 2371
QY LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
Db 2372 AAACCTGAGGANAATATCTTGAAAAGGTGGAGGCTGCATCAGGATTTACTCTGAAGAAAA 2431
QY AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2432 GATCTCTGAGAAATTTCTGGAATATTTCTGTTTCATCATATTTAAAGGTGAGAACCTTTGCTA 2491
QY LysIleArgSerIleGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2492 AAAATAGATATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATCGAA 2551
QY LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsn 733
Db 2552 AAAATAGAGAAAGTTGGCGTTCCCAATTCACAGATTGTAGATGGTCTTTTATCAAC 2611
QY SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGly 753
Db 2612 AGTAACCTGAAATTTGCAGAGGCACCATCATGTCTGATTATTCAGATGCCTCGATTGGA 2671
QY LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2672 AAAGACTTTAAACTATTTAAAAAAATTTTTCTCTCTGGAATTTAAATATAACAGATTTA 2731
QY LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2732 CTTGAAGACACTCCACAGACTGCCGATATGTGGAGGGCTTGCATGTATGATGAGTGAGA 2791
QY GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2792 GAATGCTACGACCATCCGACATCTCAGCTGGAAAAAATCAAGCAGTTTTGTAAACCTGC 2851
QY AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2852 AACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATAACCCAGTGTCACTT 2911
QY ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2912 CCAAGAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTGCCAGAAATATGGAG 2971
QY LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2972 TTATTTGCTGTTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGATATGGAG 3031
QY AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 3032 GACGATTTGCTGGTCTTTCTTTTGACAGCATGCCGATCGGATCGGTGTCAGATGGC 3091

QY 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu 913
 Db 3092 TTCAACATTCCTCAAGTCACCCCATGCCAAGAGTAGGAGACTTGAAGATGCTCTG 3151
 QY 914 GluAspLeuHisSerLeuAspSerArgGileGlnGlyCysAlaArgArgLeuLeuCys 933
 Db 3152 GAAGACCTGCATTCCTTGACTCCAGAGAGATCAAGGCTGTGCACGAAGACTGCTTTGT 3211
 QY 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 3212 CATGCATATATGTCATGTACCAAGAGTCCCAACAATGAGTTGTACAAA 3259
 RESULT 6
 CQ834226
 LOCUS CQ834226 3311 bp DNA linear PAT 29-JUL-2004
 DEFINITION Sequence 97 from Patent WO2004058805.
 ACCESSION CQ834226
 VERSION CQ834226.1 GI:50833763
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Matsuda, A. and Yoneta, S.
 TITLE T cell activating gene
 JOURNAL Patent: WO 2004058805-A 97 15-JUL-2004; -late
 Asahi Kasei Pharma Corporation (JP)
 FEATURES
 source
 1..3311
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 243..3113
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAH05328.1"
 /db_xref="GI:50833764"

CDS

/translation="MSSGLWSQEKVTSPYWEERI FYLLQECSTVDKQKLKVPKG
 SIGQYIDRSVGHRI PSAKKKNOIGLILQPHAVLFVDEKDVINEKEITELLA
 ITNCSERFLKRNRLSKGLQIDVCPVKVQLRSGEKPFGVRRFGPILAEITVSG
 IFPGVELLEEGQGGTGVYQKQLFQCDGCEVVALDKLELIEDDTALSDYAG
 PDDTQVLEPLLEINSRVSLKGETTESGTIVFCVLPKGSILGYEVGDMNDPIGNW
 DGRFGVQLCSFACVETILLINDIIIPALSESVOERPPKLAFLMRSGVGDKSSSH
 NKPKATGSDPQGNRRSELYVTLNGSSVDSQPSKSNWTYIDEVAEPKSLTEIS
 TPDFRSPLQPPVNSLTENRFSLPSLTMPNTNGSIGHSPLSLSAQSVWEELN
 TAPQESPLAMPNGSHGLEVSLAEVRENPFYGVIRWIGQPPGNEVLGLELED
 ECAGCTDGTFRGTRYETCALKALFKVLSKCRPSDFSLQPVSNQIERCNSLAFGY
 LSEVVEENTPPMEKEGLIMIGKKGIQHYNSCYLDSLTFLCLFAFSVLDTVLLRP
 KEKNVEYVSETOELARTVINPLRYGVYCATKMKLKKILEKVEAASGFTSEKDP
 BEFLNLPHHILRVEPLLRKIRAGQKQVQCYFQIEMEKNEKVGVPITQOLLEWSPIN
 SNLKEAEPSCILLIQMRPGKQFKPKIIFPSLELNLITLLEDTPQRCRIGGLAWYE
 CRECYDDPDISAGIKQFCNTQNTQVHLHPKRLNHYNPVSPKDLPLDMDWRHGC1PC
 QNMFELVICIETSHYVAFVYKGDSDSAWLFDFSDMADRDGGQNGENIPQVTPCPVEGE
 YLKMSELDLHLSDRRIQGCARRLLCDAYMCMYQSPMTSLYK"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3311
 Score: 4968.50 Matches: 947
 Percent Similarity: 99.16% Conservative: 1
 Best Local Similarity: 99.06% Mismatches: 7
 Query Match: 98.70% Indels: 1
 DB: 6 Gaps: 5

US-09-671-687A-3 (1-949) x CQ834226 (1-3311)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
 Db 243 ATGAGTTCAGGCTATGGAGCCCAAGAAAAGTCACTCCACCTACTGGGAAGCGGATT 302
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40

Db 303 TTTTACTTGTCTTCAAGAAATGCAGCGTTACAGACAAACAAACAAAGAGCTCTCTAAA 362
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
 Db 363 GTACCGAAGGAGTAGTATAGACACTATATTCAAGATCGTTCTGTGGGCATTCAGAGATT 422
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 Db 423 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAATAAATCTAGACCAACCTCATGCA 482
 QY 81 ValLeuPheValAspGlu---AspValValGluLeuLeuGluLysPheThrGluLeu 99
 Db 483 GTTCTCTTTTGTGATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACTTT 542
 QY 100 LeuAlaIleThrAsnCysGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 Db 543 TTGGCAATTACCAATTTGTAGGAGAGGTTCCAGCTCTGTTTAAACACAGAAACAGACTAAGT 602
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGlu 139
 Db 603 AAAGGCTCCAAATAGACGTGGGCTGCTGTGAAAAGTACAGCTGAGATCTGGGGAAGAA 662
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 Db 663 AAATTTCTCGAGTTGTACGCTTCAGAGAGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 722
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 Db 723 ATATTCTTTGAGTTGAATTTGCTGGAAGAGGCTGCTGCTCAAGGTTTCACTGACGGGTG 782
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 Db 783 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGCGCTGTTGTTGTCATTGGAC 842
 QY 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuLeuGluSerAspTyrAlaGlyProGly 218
 Db 843 AAGCTAGAACTCATAGAAGATGATGACACTGCTGTTGAAAGTGAATACGAGGTCCTGGG 902
 QY 219 AspThrMetGlnValGluLeuProLeuGluLeuLeuSerArgValSerLeuLysGly 238
 Db 903 GACAAATGACGCTGCAACTTCTCTTTGAAATAAACTCCAGAGTTCTTTGAAGTTT 962
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 Db 963 GGAGAAACATAGAACTCTGGAACAGTTATATTCTGTGATGTTTGTCCAGGAAAGAAAGC 1022
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 Db 1023 TTAGGATATTTTGTGTTGGTGGACATGCAACCTATTGGCAACTGGGATGGAAGATT 1082
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuHisIleAsn 297
 Db 1083 GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTGAAAGTACAATTCATTATGCACATCAAT 1142
 QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
 Db 1143 GATATCATCCAGCTTTTATCAGAGAGTGTGACGAGGAGGAGGAGGCTCCCAACTTCC 1202
 QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThr 334
 Db 1203 TTTATGTCAAGAGGTGTTGGGACAAAGGTTTCCAGTCAATAATAACCAAGAGGCTACA 1262
 QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrIleAsnGly 353
 Db 1263 GGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAAATTTATTTATACCTTAAATGG 1322
 QY 354 SerSerValAspSerGlnProGlnSerLysSerLysThrTrpTyrIleAsnGluVal 373
 Db 1323 TCTTCTGTTGACTCACAAACCAATCCAAATCAAAATAATACATGGTACATTGATGAGTT 1382
 QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393

Db 1383 GCAGAGACCCCTGCAGAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCA 1442
Qy 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
Db 1443 CCACCTCCAGCCCTCCTCTGTGAACCTCACTGACCCAGAGAACAGATCCACTCTTTTACCA 1502
Qy 414 PheSerLeuThrThyMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
Db 1503 TTCAGTCTCACCAAGATGCCCAATCAATGGAAGATTATGGCCACAGTCCACTTTCTCTG 1562
Qy 434 SerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGlnSerProProLeu 453
Db 1563 TCAGCCCACTCTGTAAATGAGAGCTTAACACTGACCCGCTCCAGAGAGTCCACCCCTTG 1622
Qy 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1623 GCCATGCCCTCTGGGAACCTCACATGCTCTAGAGTGGGCTCATTTGCTGAAGTTAAGGAG 1682
Qy 474 AsnProProPheThrGlyValIleArgTrrPleGlyGlnProProGlyLeuAsnGluVal 493
Db 1683 AACCCCTCTCTTATGGGGTAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTG 1742
Qy 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db 1743 CTCGCTGGAGCTGAACTGGAAGATGAGTGTGAGGCTGTACGATGGAACCTTTCAGAGGC 1802
Qy 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
Db 1803 ACTCGGTATTTCACCTGTGCCCTGAGAGAGGCGCTGTTGTGAACTGAGAGCTGCAGG 1862
Qy 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 1863 CCTGACTCTAGGTTTCATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTGTAACCTCT 1922
Qy 554 LeuAlaPheGlyGlyTyrLeuSerGluValValGluLeuAsnThrProProLysMetGlu 573
Db 1923 TTAGCATTTGGAGGCTACTTTAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAA 1982
Qy 574 LysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSer 593
Db 1983 AAGAAGGCTTGGAGATATGATTGGGAAGAGAAAGGCATCCAGGTCATTACATTTCT 2042
Qy 594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2043 TGTACTTAGACTCAACCTATTCTGCTATTGCTTTTGTAGTCTGTTCTGGACACTGTG 2102
Qy 614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu 633
Db 2103 TTACTTTAGACCCCAAGAAAGAACGATGTAGAAATATTATAGTGAAACCCCAAGAGCTACTG 2162
Qy 634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
Db 2163 AGACAGAAATTTGTAATCTCTGAGAAATATGGAATATGTTGTGTCACCAAAATATG 2222
Qy 654 LysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673
Db 2223 AACTGAAGAAATACTTGAAGGTGGAGGCTGCATCAGGATTTACCTCTGAGAGAAA 2282
Qy 674 AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu 693
Db 2283 GATCCTGAGGAATTTGTAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTA 2342
Qy 694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
Db 2343 AAAATAAGATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGAA 2402
Qy 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrrPheIleAsn 733
Db 2403 AAAAATGAGAAAGTTGCGCTGCCCAATTCAGAGTGTGTAGATGGCTTTTATCAAC 2462
Qy 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGly 753
Db 2463 AGTAACCTGAAATTTGCAGAGGCACCATCATGTCTGATTATTTCAGATGCTCGCTGATTGGA 2522

Qy 754 LysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2523 AAAGACTTTAAACTATTAAAAAAATTTTCTCTCTCTGGAATTAATAATATAACAGATTTA 2582
Qy 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2583 CTTGAAGACACTCCCAAGACAGTCCCGATATGTGGAGGGCTTGCATGTATGATGTAGTA 2642
Qy 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2643 GAATGCTACGACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTGC 2702
Qy 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2703 AACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATACCCAGTGTCACTT 2762
Qy 834 ProLysAspLeuProAspTrrPheArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2763 CCCAAAGACTTACCCGACTGGGACTGGAGACGCTGCATCCCTTCCAGAAATATGGAG 2822
Qy 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2823 TTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTCTTTGTGAAGATATGGGAAG 2882
Qy 874 AspAspSerAlaTrrPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 2883 GACGATTCCTCGCTGGCTCTCTTTGACGATCGCCGATCGGATGCTGTGTCAGATGGC 2942
Qy 894 PheAsnIleProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeu 913
Db 2943 TTTCAATTTCTCAAGTCACTCCCATGCCAGAAAGTAGGAGACTTTGAAGATGCTCTGT 3002
Qy 914 GluAspLeuHisSerLeuAspSerArgIleGlnGlyCysAlaArgGluLeuCys 933
Db 3003 GAAGACTGCTATTCCTTTGGACTCCAGAGAAATCAAGGCTGTGCACGAAGACTCTTGT 3062
Qy 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3063 GATGCATATATGTGCATGTACCAAGAGTCCCAACATGAGTTGTATACAA 3110

RESULT 7

LOCUS CQ719792 3480 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 5726 from Patent WO02068579.
ACCESSION CQ719792
VERSION CQ719792.1 GI:42280649
KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof

JOURNAL

Patent: WO 02068579-A 5726 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES

source

1. 3480
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3480
Score:	4962.00	Matches:	947
Percent Similarity:	99.27%	Conservative:	0
Best Local Similarity:	99.27%	Mismatches:	2
Query Match:	98.57%	Indels:	5
DB:	6	Gaps:	4

QY 716 ulysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
 Db 2395 GAAAGTGGCGTTCACCAATTCAGCAGTGTGTAGAAATGGCTTTTATCAACAGTAACT 2454
 QY 736 ulysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyAspPh 756
 Db 2455 GAAATTCAGAGGACCAATCAATGCTGATTAATTCAGATGCCCTCGATTTGGAAAGACTT 2514
 QY 756 elysLeuPheLysLysIlePheProSerLeuLeuLeuAsnIleThrAspLeuGluAs 776
 Db 2515 TAACTATTAAAAAATTTTCTCTCTGGAATTAATTAATCAAGATTTACTTGAAGA 2574
 QY 776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy 796
 Db 2575 CACTCCAGACAGTCCCGGATATGTGGAGGCTTGCATATGATGATGATGAGTATGCTA 2634
 QY 796 rAspAspProAspIleSerAlaGlyIleLysGlnPheCysLysThrCysAsnThrGl 816
 Db 2635 CGACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTGCAACACTCA 2694
 QY 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836
 Db 2695 AGTCCACCTTCATCCGAGAGGCTGAATCATTAATTAATCAACAGTGTCTACTCCCAAGA 2754
 QY 836 pleuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856
 Db 2755 CTTACCGACTGGGACTGGAGACACGGCTGCATCCCTGCCAGAAATATGGAGTATTTCG 2814
 QY 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876
 Db 2815 TGTCTCTGCATAGAAACCAAGCAGCTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTC 2874
 QY 876 rAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIl 896
 Db 2875 TGCCCTGGCTCTCTTTGAGCAGCATGCCGATCGGATGGTGGTCAAGATGGCTTCAACAT 2934
 QY 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
 Db 2935 TCCTCAAGTCACCCATGCCCAGAGTAGGAGAGTACTTGAAGATGCTCTCGAAGACCT 2994
 QY 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936
 Db 2995 GCATTCCTTGGACTCCAGGAGAAATCAAGGCTGTGCACGAGACTGCTTGTGATGCATA 3054
 QY 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 3055 TATGTGCATGTACAGAGTCCCAATGATGATTTGTACAA 3094

RESULT 8
 BD231207
 LOCUS 4527 bp DNA linear PAT 17-JUL-2003
 DEFINITION Human cytoskeleton associated proteins.
 ACCESSION BD231207
 VERSION BD231207.1 GI:33040977
 KEYWORDS JP 2002526076-A/9.
 SOURCE Homo sapiens (human)

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 4527)

REFERENCE
 AUTHORS Lal P., Tang, T.Y., Yue, H., Hillman, J.L., Bandman, O., Corley, N.C.,
 Guegler, K.J., Patterson, C., Azimzai, Y. and Baughn, M.R.
 Human cytoskeleton associated proteins
 Patent: JP 2002526076-A 9 20-AUG-2002;
 INCYTE PHARMACEUTICALS INC

TITLE
 JOURNAL
 COMMENT OS Homo sapiens (human)
 PN JP 2002526076-A/9
 FD 20-AUG-2002

PF 17-SEP-1999 JP 2000574254
 PR 18-SEP-1998 US 60/172226, 27-APR-1999 US 60/131321 PI
 PREETI LAL, TOM Y TANG, HENRY YUE, JENNIFER L HILLMAN, OLGA PI
 BANDMAN,

PI NEIL C CORLEY, KARL J GUEGLER, CHANDRA PATTERSON, YALDA AZIMZAI,
 PI MARIAH R BAUGHN
 PC C12N15/09, A61K38/00, A61K45/00, A61P1/16, A61P3/00, A61P3/06 PC
 , A61P3/10, A61P5/00,
 PC A61P5/38, A61P7/00, A61P7/06, A61P9/00, A61P9/10, A61P11/06 PC
 , A61P13/00, A61P13/12
 PC A61P17/06, A61P19/02, A61P19/10, A61P21/04, A61P25/00, A61P25/00,
 PC A61P25/04,
 PC A61P25/14, A61P25/16, A61P25/28, A61P29/00, A61P31/04, A61P31/12,
 PC A61P31/18,
 PC A61P33/00, A61P35/00, A61P35/02, A61P37/00, A61P43/00, C07K14/47,
 PC C07K15/18,
 PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12P21/08 PC
 , C12Q1/68, G01N33/15,
 PC G01N33/50, G01N33/53, G01N33/53, G01N33/56, C12N15/00, C12N5/00,
 PC A61K37/02
 CC Incyte ID No: 2363327
 FH Key Location/Qualifiers
 FT source 1. .4527
 FT /organism='Homo sapiens (human)'.
 FT /db_xref='taxon:9606'.
 FEATURES
 source
 1. .4527
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 4527
 Score: 4942.00 Matches: 944
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 96.17% Indels: 5
 DB: 6 Gaps: 4
 US-09-671-687A-3 (1-949) x BD231207 (1-4527)

QY 4 GlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIlePheTyrIeu 23
 Db 270 GGCTTATGAGCCCAAGAAAGTCACTTCACTCCCTACTGGGAAGAGCGGATTTTACTTG 329
 QY 24 LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLysValProLys 43
 Db 330 CTTCTTCAAGATGACGCGTTACAGACAAACAAAGCTCTCTTAAAGTACCGAA- 388
 QY 44 GlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIleProSerAla 63
 Db 389 GGAAGTATAGACAGTATATTCAGATCGTTCGTGGGGCATTCAGGATTCCTTCTGCA 448
 QY 64 LysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAlaValLeuPhe 83
 Db 449 AAAGGCAAGNAATCAGATTGATTAATAATTCAGACCACTCATGCGATTCCTCTT 508
 QY 84 ValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeuAlaIle 102
 Db 509 GTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTTTTGGCAAT 568
 QY 103 ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgenArgLeuSerLysGlyLeu 122
 Db 569 ACCAATGTGAGAGAGGTTTCAGCTCTTTTAAACAGAACAGACTAAGTAAGGAGGCTC 628
 QY 123 GlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGluLysPhePro 142
 Db 629 CAATAGACCTGGGCTGCTCTGTGAAGTACAGCTGAGATCTGGGGAGAGAAATTCCT 688
 QY 143 GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe 162
 Db 689 GGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGGACAGCTCCGGAATATCTTT 748
 QY 163 GlyValGlnLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGly 182
 Db 749 GGAGTTGAATGTGGGAAGAGGTCGTGGTCAAGGTTTCTACTGACGGGGGTGTACCAAGG 808

Db 2969 ACCCATGCCCAAGTAGGAGAGTACTTGAAGATGCTCTCGAAGACCTGCAATTCCTTG 3028
 Qy 920 AspSerArgArgileGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMet 939
 Db 3029 GACTCCAGAGAAATCCAGGCTGTGCAGAAAGCTCTTTGTGATGCATATATGTGCATG 3088

Qy 940 TyrGlnSerProThrMetSerLeuTyrLys 949
 Db 3089 TACCAGAGTCCAAATGAGTTGTGTACAAA 3118

RESULT 9
 LOCUS AKI22389 4314 bp mRNA linear ROD 15-MAR-2003
 DEFINITION Mus musculus mRNA for mKIAA0849 protein.
 ACCESSION AKI22389

VERSION AKI22389.1 GI:28972434

KEYWORDS FLI_CDNA.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE
 AUTHORS Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Aizawa, H., Yuasa, S., Nakajima, D., Nagase, T., Ohara, O. and Koga, H.

TITLE Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries

JOURNAL DNA Res. 10, 35-48 (2003)

REFERENCE
 AUTHORS Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.

TITLE Direct Submission

JOURNAL Submitted (07-FEB-2003) Hiasashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics, 2-6-7

Kazusa-kamata, Kisarazu, Chiba 292-0818, Japan
 (E-mail: mousekazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)

COMMENT The CRATE program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.

FEATURES
 Location/Qualifiers
 1. 4314

/organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mbh04492"
 /tissue_type="brain"
 /dev_stage="adult"
 /notes="vector:modified pBC SK+"

1. 4314
 /gene="mKIAA0849"
 <197..3196

/gene="mKIAA0849"
 /notes="CDS is predicted by in silico analysis. Start codon is not identified."

/codon_start=1
 /evidence=not_experimental
 /product="mKIAA0849 protein"
 /protein_id="BAC65671.1"
 /db_xref="GI:28972435"

translation="IAVFLQHGCOVAKILLMDLQVRHFFGDRFTSGSFLPVNVLKV
 TVTMSSGLSQEAKVTSFYHEERIFXLLQECSTVDKOTOKLLKVPKSGIGXVIQDRSV
 GHSRVSPTGGKQNGIKLEQPHAVLFVDEKDVINEKFEIALLAIINICERLSLF
 RNRRLSKGLQVGVSPVQVRSGEEKFPVGVFRGPLLAERTVSGIFVGLLEEG
 RGQGTDTGVYQKLPQCDCEGDFVALDKLEIEDDNGLESDFAGPGDMQVPPPP
 LEINRSVLKVGSESTGTVICDVLPGKESLGYFVGVMDNDPNIQWMDGRFDGVQLCS
 FASVESTILLHINDIIPDSVTQERRPKLAFMSRGVGDKSSSHNPKVTGSDSPGS
 RNRSELFTYLLNSVDSQSKSKPNYIDEVAEDPAKSLTEMSDDPHSSPPQPPSM
 NLSSENRRHSLFSLTKMPTNGNSMAHSPLSLVSQVMSGLNSTPVSQSPPLFISG
 NAHLGVSLAEVKNPPFVGIWIGQPPGLSDVLAGELEDECACTDGTGTRTRY
 FTCLAKLKFALVKRCSRFASLQPNQIERCNLSLAPGGYLSVEVENTPPRMEK
 EGLEIMIGKKGIQGHNSCYLDTLFCLPAPSSALDVLRLRPKNDIEYSETQEL
 LRTEIVNPLRIYGVVCAATKIMLKILEKVEAASGFTSBEKDPPEEFNLINILPHDLRAVE

gene
 CDS

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 4314
 Score: 4718.50 Matches: 898
 Percent Similarity: 96.54% Conservative: 22
 Best Local Similarity: 94.23% Mismatches: 28
 Query Match: 93.73% Indels: 5
 DB: 10 Gaps: 5

US-09-671-687A-3 (1-949) x AKI22389 (1-4314)

Qy 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
 Db 338 ATGAGTTTCAGGCTGTGGAGCCCAAGAGAAAGTTTACTTCACTCCCTACTGGGAAGACCGGATT 397
 Qy 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 Db 398 TTTTATCTCTCTTCTCAAGATGCAGTGTAAACACAACTCAGAAAGCTGCTGAAA 457
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
 Db 458 GTACCCAAAGGAGTATAGACAGTACATCCAAAGACCGTTCTGTGGGCATTCAGAGTT 517
 Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 Db 518 CCTTCCACAAAAGGCAAGAAAATCAGATTGGATTAAAAATCTTGGAGCAACCGCATGCA 577
 Qy 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeuLeu 99
 Db 578 GTTCTCTTTGTGATGAAAAGGATGTTGTAGAAATAAATAAATAATTCACAGAGTTACTG 637
 Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 Db 638 TTGGCAATTACCACTGTGAGGAGAGGCTCAGGCTATTATAGAACAGACTCCGACTAAGT 697
 Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 Db 698 AAAGGCTCCAGTACAGCTGGGAGCTCTCTGTGAAAAGTACAGCTGCATCTGGGGAAGAG 757
 Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 Db 758 AAATTTCCAGGAGTTGTACGCTTCAGAGGACCTTTATTAGCGAGAGACGCTGTCGGGG 817
 Qy 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 Db 818 ATTTTCTTTGGAGTAGAATTTTGGAGAAGGTCGTGCTCAGGTTTCACGATGGGGTA 877
 Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 Db 878 TATCAAGGGAAGCAGCTTTTCCAGTGTGTAGAGACTGTGCGCGTGTGTTGTCATTGGAC 937
 Qy 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 Db 938 AAACCTAGAACTTATAGAAATGATGACAAATGGATTGGAAATGATTGTCAGGCCCAAGGA 997
 Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 Db 998 GATACAAATGAGTTGTAACCTCCCTCTTTGGAATAATAACTCCAGAGTTTCTTGAAGTT 1057
 Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 Db 1058 GGAGAAAGTACAGAACTCTGGAACAGTAATATCTGTGATGTTTACCAGGAAAAAGAGACT 1117
 Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnThrAspGlyArgPhe 278
 Db 1118 CTAGGATATTGTTGTTGGTGGAGCATGATGATAACCTATTGGCAACTGGGATGGAGGTTT 1177

QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuHisIleAsn 297.
DB 1178 GATGGAGTACAGCTCTGTAGTTTGAAGTGTGAAGTACAAATCTCTGGCACATCAAT 1237
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 1238 GACATCATCCAGATAGCTGACACAGGAAGAGGAGCCCTCCCAACTTGGCTTTATGTCA 1297
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1298 AGAGGTGTAGGTGACAAAGGTTTCATCTAGTCATAATAAACCAAGGTTACAGGATCTACC 1357
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheThrThrLeuAsnGlySerSerVal 356
DB 1358 TCAGACCCCTGGAAGTGAAGACAGATCTGAATATATTTATACCTTAATGGGTCTCTGTT 1417
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
DB 1418 GACTCACAA---CAATCCAGTCCCAAAATCCATGGTACATTGATGAAGTTGCAGAGAC 1474
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1475 CTTGCAAAAGTCACTTACAGAGATGCTCTCGGACTTCGGACATTCATCTCTCTCCACCGCAG 1534
QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1535 CCTCTTCCATGAACCTCTGTCTAGCGAGAACAGATTCACCTCTTACCTCTTACGCTG 1594
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuAlaGln 436
DB 1595 ACAAGATGCCCAATACTAATGGCAGCATGGCTCATAGTCCACTCTCTGTGTCAGTGCAG 1654
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1655 TCTGTGATGGGGAGCTGACACAGCACCTGTCCAGAGAGTCCACCTTGGCCATCTCT 1714
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1715 TCTGGGAATGCCACACGGCTAGAGGTGGCTCACTGGCTGAAGTAAAGAGAACCCCGC 1774
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1775 TTCTATGGGGTTATCCGTTGGATGGCCAGCCACCGAGGCTCAGTACGCTGCTAGCTGA 1834
QY 497 LeuGluLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1835 CTGGAACTGGAGATGAATCGCAGGCTGTACAGATGGAACTTTCAGGGGCGCCGGTAT 1894
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1895 TTCACGTGTGCCCTGAAGAGGCACCTGTTGTGAACCTGAAGAGCTGCAGACCGGACTCT 1954
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1955 AGGTTTGCATCTTGCAGCCTGTTTCCAATCAGATTGAAAGGTGTAACTCTTTAGCATTT 2014
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
DB 2015 GGGGGCTATTAAAGTGAAGTAGTAGAAGAAAATACCTCCACTAAATGGAAGAGGAGGT 2074
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 2075 TTAGATATATGTTGAAGAGAGAGGATCCAGGGCCATTACAAATCTTGTGTACTTA 2134
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2135 GACTCACTTATCTGCTTATTTGCTTTTGTCTGCTGCTGACACTGTGTACTTAGA 2194
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuArgThrGlu 636
DB 2195 CCCAAAGAGAGAAATGATATAGATATTACAGTGAGACTCAGGAGCTACTGAGGACAGAG 2254
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656

DB 2255 ATAGTCAATCTCTGAGATATATGATATGTGTGCCCAAAAGATTATGAACCTGAGG 2314
QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 2315 AAAATACTTGAAGAGTTGAGGCTGCATCAGGTTTACCTCTGAAGAAAAAGATCTCGAA 2374
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2375 GAATCTTAAATATCTCTGTTTCATGATATTAAAGGGTTGAACCATTTGTTAAAAATAAGA 2434
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2435 TCAGCAGGTCAAAAGTTCAAGACTGTAATCTTATCAAAATTTTATGGAAAAAATGAG 2494
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2495 AAAGTTGGAGTACCCCAATTCAGAGTTATTAGATGGTCTTTTATCAACAGCAACCTG 2554
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
DB 2555 AAATTTGCAGAGGCACCATCATCTTGAATTTATCCAGATGCTCGGTTTGGAAAAAGACTTT 2614
QY 757 LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2615 AAATATTTAAAAAAATTTTCTCTCCCTGGAATTAATAATATAACAGATTACTTTGAAGAC 2674
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2675 ACTCCAGCAGGAGTCCGCATCTGTGGAGACTTGCAGTGTACGAGTGCAGGGAGTCTAT 2734
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2735 GACGATCCGAGACTCTCAGCTGGGAAGATCAAGCAGTTCTGTGAAGACTGCAGCACTCAG 2794
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2795 GTTCACCTTCATCCAGGAGGTTGAATCATCTTATCATCCAGTATCATCTCCCAAGAC 2854
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 2855 TTCCCTGACTGGGACTGGAGACACGGCTGCATCCCTGTGAGAGATGGAGTTATTGCT 2914
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
DB 2915 GTTCTCTGCATAGAACGACCACTATGTTGCTTTTGTGAAGTATGGAGGATGACTCT 2974
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 2975 GCCTGGCTTTCTTTGACAGCATGGGAGATCGAGATGGTGGTTCAGAAATGGCTTCAACATT 3034
QY 897 ProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeu 916
DB 3035 CCACAGTGCAGCCCTGCCAGAGTGGGAGAGTACTTGAAGATGTCTCTGGAGGACCTG 3094
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
DB 3095 CACTCTTTGGACTCCAGAGGATTCAGGCTGTGGCGGAGACTTCTTTGCGATGCATAC 3154
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 3155 ATGTGCATGTACCAGAGTCCCAACCATGAGCTGTACAAA 3193
RESULT 10
BC042438
LOCUS
DEFINITION
Mus musculus cyclinomatosis (turban tumor syndrome), mRNA (cdna
clone MGC:25429 IMAGE:3983771), complete cds.
ACCESSION
BC042438
VERSION
BC042438.1
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE AUTHORS

1 (bases 1 to 4501)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viallon, D.K., Muzny, D.B., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shuevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 4501)
Strausberg, R.

Direct Submission
Submitted (02-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILML)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseg, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILML at: <http://image.llnl.gov>
Series: IRAK plate: 30 Row: g Column: 3.

FEATURES

source

Location/Qualifiers
1..4501
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:25429 IMAGE:3983771"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP Mam1"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..4501
/gene="Cylid"
/notes="synonyms: EAC, CDMT, CYLDL, mKIAA0849"
/db_xref="LocusID:74256"
/db_xref="MGI:1921506"
176..3034
/gene="Cylid"
/codon_start=1
/product="Cylindromatosis (turban tumor syndrome)"
/protein_id="AAH42438.1"
/db_xref="GI:27503671"

gene

176..3034
/gene="Cylid"
/codon_start=1
/product="Cylindromatosis (turban tumor syndrome)"
/protein_id="AAH42438.1"
/db_xref="GI:27503671"

CDS

176..3034
/gene="Cylid"
/codon_start=1
/product="Cylindromatosis (turban tumor syndrome)"
/protein_id="AAH42438.1"
/db_xref="GI:27503671"

/db_xref="LocusID:74256"
/db_xref="MGI:1921506"

/translation="MSSGLSQEKTSPYWEERIFYLLQECSTVDKQTKLLKPKVG
STGQIQDRSVGHRSVPSTKGNQIGLILQPHAVLFVDEKDVVVEINEKFTLELLA
ITCERLSLFRNRLSLKGLQVGVSPKVLORSGEERKPGVVRFRGPLLAERTTVSG
IFGVELLEEGRGQGTGQVYOGKOLFOCDEDCGVFEALDKLELIEDDNGLESDFAG
PDGTWQVPPPLEINSRSLKVGSTESGTIVFCVLPKESLGIVYGVGDMNDPIGNW
DRFGVQLCSFASVESTILLHNDIIPDSVTOERPPPKLAFMRSGVGTSSSHNRP
KVTGTPDPSGSRNRSLEFYTLNGSSVDSQSKSNKWPYIDEVAEDPAKSTENSSDFG
HSSPPPLPSSNNLSSENFSLFSLTPMNTNGSMASPLSLVSQVNGELNFTSPV
QSSPPLTSSGNAHGLVGLAEVKNPPFYGVIRWIGOPGLSDVLAGELEDECAV
CTDGTGRTVETKALKFVKLSCRPSRPFASLOPVSNQIERNCSLAFGGVLSV
VEENTPPKKEGLEIMIGKKGIQGHYNSCHLDSTLFCFLFAFSSALDVTLLRPKERN
DIEYFSETOELLRTIENPLRIGYVCAIKMKRLKLEKVEAASGTSEKPOPEEFL
NIFHDIILVEPLLRISAGQVQDCNFYQIFMEKNKVGVTIQLLEWSFTNSNLK
FAEAPSLTIOMPRFKDPKLPKIPPSLELNTDLLEDTPROCRICGGLAMTECRRC
YDDPDISAGIKQFCCKTSTOVHLHPRRLNHSVHPVSLPKDLDPMDWRHGCICOKME
LPAVLCIETSHYVAFVKYKDDSAWLPFDSMADRDGQNGFNIPQVTPCEVGEYLMK
SLEHLSLDSRRRIQCARLLCDAYNCMTQSPMTSLYK"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 4501
Score: 4718.50 Matches: 898
Percent Similarity: 96.54% Conservative: 22
Best Local Similarity: 94.23% Mismatches: 28
Query Match: 93.73% Indels: 5
DB: 10 Gaps: 5

US-09-671-687a-3 (1-949) x BC042438 (1-4501)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
DB 176 ATGAGTTTCAGGCTGTGGAGCCAGAGAAAGTTTACTTCACTTCTGGAGAACGAGATT 235

QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 236 TTTTATCTCTCTTCTCAAGAATGTCAGTGTAAACAGACAAACAACTTCAGAGCTGCTGAAA 295

QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
DB 296 GTACCCAAAGGAGTATAGACAGTACATCCAGACCGTCTCTGTGGGCGATTCAAGAGTT 355

QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 356 CCTTCCACAAAAGGCAAGAAAATCAGATTGGATTAAAAATCTTGGAGCAACCGCATGCA 415

QY 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGlnLysPheThrGluLeuLeu 99
DB 416 GTTCTCTTTTGTGATGAAAAGGATGTTGTAGAAATAAATGAAAATTCACAGAGTTACTG 475

QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 476 TTGGCAATTACCAACTGTGAGGAGAGCTCAGCTTATTAGAAACAGACTCCGACTAAGT 535

QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 536 AAAGGCTCCAGGTAGCGTGGCGAGTCTCTGTGAAAGTACAGCTGCGATCTGGGGAAGAG 595

QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 596 AAATTTCCAGGAGTGTGACGCTTCAGAGGACCTTTATTAGCGGAGAGACGCTGTCGGGG 655

QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
DB 656 ATTTTCTTTTGGAGTAGAATTATTGGAGAAAGGTCGTCGTCAAGTTTTCACGAGTGGGTA 715

QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 716 TATCAAGGGAAGCAGCTTTTCCAGTGTGATGAGGACTGTGGCGGTGTTTGTTCATTTGAC 775

QY 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 1999 LysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218

Db 776 AACTAGAACTTATAGACATGATGACAATGGATTGGAAGTGATTTTCAGGCCCCAGGA 835
Qy 219 AspThrMetGlnValGluLeuProLeuGluLeuAAsnSerArgValSerLeuLysGly 238
Db 836 GATACAAATGACAGTTGAACCTCCCTTTTGGAAATAAATCCAGAGCTTTCTTTGAAGGTT 895
Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 896 GGAGAAAGTACAGAAATCTCGAAGCAGTAATCTCTGTGATGTTTACAGGAAAGAGAGT 955
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 956 CTAGGATATTTGTTGTTGAGATGAGATAACCTATTGGCAACTGGGATGGAAGGTTT 1015
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1016 GATGGAGTACAGCTCTGTAGTTTTCAGAGTGTGAAAGTACAAATCTCTCGACATCAAT 1075
Qy 298 AspileleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
Db 1076 GACATCATCCAGATAGCTGACAGGAAAGAGGCGCTCCCAAACTTGCCTTTATGTCA 1135
Qy 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1136 AGAGGTGAGGTGACAAAGGTTCTCTAGTCATAATAACCAAGGTTACAGGATCTACC 1195
Qy 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1196 TCAGACCTCGGAAGTAGAACAGATCTGAATATTTTATACCTTAAATGGTCTCATCTGT 1255
Qy 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
Db 1256 GACTCACAA---CAATCCCAAGTCCAAAAATCCATGGTACATTTGATGAAGTTGCAGAGAC 1312
Qy 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
Db 1313 CTGCAAGTCACTTACAGAGATGTTCTCGAGCTTCGACATTCATCTCTCCACCGCAG 1372
Qy 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1373 CCTCTTCCATGNACTCCTTGCTAGCGAGAACAGATTCACCTCCTTACCTTCAGCCTG 1432
Qy 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLysSerLeuAlaGln 436
Db 1433 ACAAGATGCCAAATACTAATGACGATCGCTCATAGTCCACTCTCTCTGTGAGTGAC 1492
Qy 437 SerValMetGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1493 TCTGTATGGGGAGCTGAACACACACCTGTCCAGAGAGTCCACCTTGGCCCATCTCT 1552
Qy 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1553 TCTGGGAATGCACACGGCTAGAGTGGGCTCACTGGCTGAATGAAGAAACCCCGC 1612
Qy 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1613 TTCTATGGGTTATCCGTTGGATTGGCCAGCCACCGGCTCAGTGAGTGCTAGTCTGA 1672
Qy 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1673 CTGGAACCTGGAAGATGAATGCGCAGGCTGTACAGATGGAACTTTTCAGGGGCGACGCGTAT 1732
Qy 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1733 TTCACGTGTGCCCTGGAAGAGGACCTGTTTGTGAACCTGAAGAGCTGACAGCCGACTCT 1792
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1793 AGGTTTGCATCCTTGACGCTGTTTCCATCAATGAAGGTGTAACCTCTTTAGCATTT 1852
Qy 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1853 CGGGGCTATTTAAGTAGAGTAGTAGAAGAAATACTCCACCTTAAATGGAAAGGAAGGT 1912

Qy 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCystTyrLeu 596
Db 1913 TTAGAGATAATGATTTGGAAAGAGAAAGGATCCAGGGCCCATTAACAATCTTGTGTACTTA 1972
Qy 597 AspSerThrLeuLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 1973 GACTCAACTTATCTGCTTATTTGCTTTTAGTTCTGCTCCCTGGACACTGTGTACTTAGA 2032
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2033 CCCAAAGAGAAGAAATGATATAGAGTATTCAGTGCAGACTCAGGAGCTACTCAGGACACAG 2092
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2093 ATAGTCAATCCTCTGGAATATATGATATGTGTGTGCCACAAAAGATTATGAACATGAGG 2152
Qy 657 LysIleLeuGluLysValGluAlaAspGlyPheThrSerGluGluLysAspProGlu 676
Db 2153 AAAATACTTGAAGAGTTGAGGCTGCATCAGTTTTTACCTCTGAAGAAAAGATCCTGAA 2212
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2213 GAATCTTAAATATCCTGTTTCATGATATTTTAAGSGTTGAACCATTCGTTAAAATAAGA 2272
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2273 TCAGCAGGTCAAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGGAAAAAATGAG 2332
Qy 717 LysValGlyValProThrIleGlnGlnLeuLeuLutrpSerPheIleAsnSerAsnLeu 736
Db 2333 AAAGTTGGAGTACCCAAATTCAGCAGTATTAAGATGGTCTTTTATCAACAGCAACCTG 2392
Qy 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
Db 2393 AAATTTGCAGAGGCACCATCATGCTTGATTTATCCAGATGCCTCCGTTTGGAAAGACTTT 2452
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2453 AAACATTTTAAAAAATTTTCTCTCCCTGGAATTAATAATAACAGATTTACTTTGAAGAC 2512
Qy 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2513 ACTCCCGAGGAGTCCCGCATCTGTGGAGGACTTCGGATGTACGAGTGCGAGGAGTGCTAT 2572
Qy 797 AspAspProAspIleSerAlaGlyLysGlnPheCysLysThrCysAsnThrGln 816
Db 2573 GACATCCGACATCTCAGCTGGGAAGATCAAGCAGTTCGTGAAGCTTGAGCACTCAG 2632
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2633 GTTCACCTTCATCCAGAGGTTGAATCATTTTATCATCCAGTATCACTTCCCAAGAC 2692
Qy 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2693 TTGCCTGACTGGGACTGGACACGCGTGCATCCCTGTGCAGAAAGATGGAGTATTTTCT 2752
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2753 GTTCTCTGCATAGAAACGACCATATGTTGCTTTTGTGAAGTATGGGAAGATGACTCT 2812
Qy 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2813 GCCTGGCTTTCTTTGACAGCATGGCGGATCGAGATGGTGGTCAAGATGGCTTCAACATT 2872
Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2873 CCACAGTGACGCGCTGCCCAAGATGGGAGTACTTTGAAGATGCTCTCTCGAGGACCTG 2932
Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2933 CACTCTTTGGATTCAGAAAGGATTCAGGGCTGTGCGGCAGACTTCTTTTGGATGCATAC 2992

```

QY 937 MetCysMetTyTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2993 ATGTGATGTACAGAGTCCACCATGAGCTGTACAAA 3031

RESULT 11
BC082001 3267 bp mRNA linear ROD 02-SEP-2004
LOCUS Rattus norvegicus cdna clone MGC:94236 IMAGE:7131198, complete cds.
DEFINITION BC082001
ACCESSION BC082001
VERSION BC082001.1 GI:51858716
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 3267)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahay,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,K.,
Boutfard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3267)
Director MGC Project.
Direct Submission
Submitted (01-SEP-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate; 194 Row; J Column; 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1. 3267
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:94236 IMAGE:7131198"
/tissue_type="Testis, rat (Brown Norway)"
/clone_lib="NIH_MGC_238"

FEATURES
source

```

```

/lab_host="DH10B"
/note="Vector: pExpress1"
196. 3057
/codon_start=1
/product="Unknown (protein for MGC:94236)"
/protein_id="AAH82001.1"
/db_xref="GI:51858717"
/translation="MSSGLWNEKVTSPYWEERLFYLLQECSTVDKQTKLRVPKG
STGQYIDRSVGHSPVSAKKNQIGLKILEQHAFLVDEKDVVEINEKFTLELLA
ITNCEERLSLPRNIRLSKGLVDVGVSPVRVLRGPGVVRPRGPIILABRTVSG
IEFFVELLEEGRGQGTGVDGYQKQLFQCDGCGVFVLDKLELIEDDDNGLSDFAG
PGDTVQVPEPPLEINSRSLKVGSESTGVTFCDVLPKGLSGLGFVGVDMDFNPGW
DRGFQVQLCSFASVESTVLLHNDIIPDSVTQERRPPKLAFMRSRGVGVGGSSHNKP
GHSSTDPGSRNRSSELYFLANGSSVDQQSKNPNWYDEVAEDPAKSLTMSDDP
KSSPPQPPMNSLSSSENRHSLPFLSKPMPTNGSMASPLSLVSQVVGELNSTP
VOESPMPSSSGNAHGLEVAENKPPFYGVIRWIGOPPGGLSDVLGLEDECA
GCTDGTFRGTRYTCALKALFKLKSCHPDSRPASLQPVNSQIERCNSLAFGLYSE
VVEENTPPKWEKEGLIMIGKKGIQIHNYSYDSTLFCIFAFSSALDVLVRPKP
NDVEYSETOELLRTIENVLRIGYVCATKMKRRIKLEKVAASGDTSEEDKPEP
LNILFHDLRVEPLKIRAGQKQVDCNFYQIFEMKNEKVGVTIQQLLEWSTNSP
KFAEAPDCLIIQMPRGDKPLFKIPPSLELNITLLEDTPRQCRICGGLAMYECRE
CYDDPDISAGIKQFCCKTCTOVHLHPRLNHTYHPVSLPKDLDMDMRHGCI PCQKM
ELFVLCIETSHYVAFVKYKDDSAWLFPDSMADRGQGNGFNPQVTPCFVEGYLK
MSLEUHLSDSRRIQGCARRLLCDATCMYQSTMSLYK"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3267
Score: 4712.00 Matches: 893
Percent Similarity: 96.54% Conservative: 27
Best Local Similarity: 93.70% Mismatches: 29
Query Match: 93.60% Indels: 4
DB: 10 Gaps: 4
US-09-671-687A-3 (1-949) x BC082001 (1-3267)
QY 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTyrGluGluArgIle 20
Db 196 ATGAGTTTCAGGCTGTGGAACCAAGAGAAAGTTTACTTCCCTACTCGGAAGACGGCTT 255
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 256 TTTTATCTGCTTCTTCAAGAAATGCAGTGTAAACAGACAAACAGACAGACTCTCTGAGA 315
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 316 GTACCCAAAGGGAGCATAGACAGTACATCAAGACCGTTCCTGGGGGCATTCAGAGTT 375
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 376 CCTTCTGCTAAAGCAAGAAAAATCAGATTGGATTAAAAATCTTAGAGCAACCGCATGCA 435
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
Db 436 GTTCTGTGTTGATGAAAAGGATGTTGTGAAAAATAATGAAAAATTCACAGAGTTACTG 495
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgSer 119
Db 496 TTGGCAATTACCAACTGTGAGGAGAGCGCTCAGCCCTATTATGAAACAGAAATCCGACTAAGT 555
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 556 AAAGGCTTCAGGTAGACGTGGGCGAGTCTCTGTGAGAGTACAGCTCGCATCTGGGGAGAG 615
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 616 AAGTTTCCAGGAGTTGTACGCTTCAGAGGACCTTTATTAGCGGAGAGACGGTGTGCGGG 675
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlnPheThrAspGlyVal 179
Db 676 ATTTTCTTTTGGAGTAGAATTACTGGAAGAGAGTCTGCGCCCAAGTTTTCACATGATGGGTG 735
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198

```


Db	736	TATCAGGAAACACAGCTCTTCCAGTGTGATCAGACCTGGCGCTTTTGTGTGATTTGGAC	795
Qy	199	LysLeuGluLeuIleGluAepAepAepThrAlaLeuGluSerAepTyrAlaGlyProGly	218
Db	796	AAAGCTGGAGCTTATAGAAAGATGATCAATGGGTTCGAAAGTGATTTTGCAGGCCAGGA	855
Qy	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238
Db	856	GATACAGTCCAGGTGAACCTCCCTTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT	915
Qy	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258
Db	916	GGAGAAAGTACAGAACTCGAACAGTGATATCTCTGATGTTTACAGGAAAAAGAGAT	975
Qy	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAepGlyArgPhe	278
Db	976	CTAGGATATTTGTTGGTGGACATGGATTAACCTATTGGCAACTGGGACGGAAGGTTT	1035
Qy	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297
Db	1036	GATGGGTTTCAGCTTTGCAAGTTTGTCAAGTGTGGAGATGACATCTCTCTACATCAAT	1095
Qy	298	AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer	317
Db	1096	GACATCATCCAGATAGCGTGACACAGGAAGAGACCTCCCAAACTTGCCTTTATGTCA	1155
Qy	318	ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr	337
Db	1156	ACAGGTGTAGTGACAAAAGTTTCATCTAGTCATAATAACCAAAAGGTTACAGGATCTACC	1215
Qy	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356
Db	1216	TCAGACCTCGAAGTAGAAACAGATCTGAATATATTTATACCTTAAATGGGTCTCTGTT	1275
Qy	357	AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp	376
Db	1276	GACTCACAAACAATCCAAAGTCCAAAACCCCAATGGTATGATGAAGTTGCAGAGAC	1335
Qy	377	ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln	396
Db	1336	CTGCAAAAGTCACTTACAGAGATGCTTCAGACTTCGGACATTCAGCTCCACCGCAA	1395
Qy	397	ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu	416
Db	1396	CCTCCTTCATGAACCTCTGTCTAGCGAGACAGATTCACCTCTTACCTTCAGCCTG	1455
Qy	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436
Db	1456	ACAAAGATGCCCAACATAATGCGCATGGCTCACAGTCCACTCTCTCTGTGAGTGCAG	1515
Qy	437	SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro	456
Db	1516	TCTGTGATGGGAGCTGAACAGCATCTCTGTCTCAGGAGATCCACCATGCGCCAGCTCT	1575
Qy	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476
Db	1576	TCCTGGGAATGCACCGGCTAGAGTGGCTCACTGGCTGAAGTAAAGAGAACCCCGG	1635
Qy	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly	496
Db	1636	TTCTATGGGTTATCGTTGGATTGGCCAGCACCGGCTCAGTGCAGTGTCTGTCTGGA	1695
Qy	497	LeuGluLeuLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516
Db	1696	TTGGAACTGGAAAGTGAATGCGCAGGTTGTACGGATGGAACTTTCAGGGGCGACGCTAT	1755
Qy	517	PheThrCysAlaLeuLysIleAlaLeuPheValLysLeuLysSerCysArgProAspSer	536
Db	1756	TTCACTGTGGCTTGAAGAAAGACACTGTTCTGTGAAACTGAAGAGCTGCAGACCACTCT	1815
Qy	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556
Db		TATCAGGAAACACAGCTCTTCCAGTGTGATCAGACCTGGCGCTTTTGTGTGATTTGGAC	795
Qy	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly	576
Db	1876	GGGGGCTACTTAAGTGAAGTAGTAGAAGAAAATACGCCACCTAAATTCGAAAAGGAGGT	1935
Qy	577	LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596
Db	1936	TTAGAGATAATGATTTGAAAAGAAAGAGGCATCCAGGGCCATTACAAATCTTGTGTACTTA	1995
Qy	597	AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg	616
Db	1996	GACTCAACTTTATCTGCTTATTTGCTTTTAGTTCTGCCCTGGACACTGTATTAATTAGA	2055
Qy	617	ProLysGluLysAsnAepValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
Db	2056	CCCAAGAGAGAAGATGATGTAGAGTATTACAGTGAGACTCAAGAGCTACTCAGGACAGAG	2115
Qy	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656
Db	2116	ATAGTCAATCTCTGAGAATATATGGATATGTGTGTGCCCAAAAGATTTATGAAGCTGAGG	2175
Qy	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
Db	2176	AAATACTTGAAGAAAGTTGAGGCTGCATCAGATTTACCTCTGAGGAAAAAGATCCTGAA	2235
Qy	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696
Db	2236	GAATTTCTAAACATCTGTTTTCATGATATTTTAAAGGTTGAACCATTTGTTAAAAATAAGG	2295
Qy	697	SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu	716
Db	2296	TCAGCAGGTCAAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGGAAAAAATCAG	2355
Qy	717	LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu	736
Db	2356	AAAGTCGGAGTACCCCAATCCAGCAGTATTAGATGGTCTTTTATCAACAGCAACCTG	2415
Qy	737	LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe	756
Db	2416	AAATTTGCAGAGGCACCATCATGCTTGATTATCCAGATGCTCGGTTTGGGAAGACTTT	2475
Qy	757	LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp	776
Db	2476	AAACTATTTAAAAAATTTTCTCTCCCTGGAATTAATATAACAGATTTACTTTGAAGAC	2535
Qy	777	ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr	796
Db	2536	ACTCCAGGCAGTGCCTCATCTGTGGAGGACTCGCCATGTATGATGTGTAGAGATGCTTAT	2595
Qy	797	AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln	816
Db	2596	GATGACCCGACATCTCGGACGGAGAGATCAGCAGTTCTGTAAAGCTTGAGACTGAGACTCAG	2655
Qy	817	ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp	836
Db	2656	GTTCACTTTCATCCAGAGACTGAATCACACTTACCATCCAGTATCACTTCCCAAGAC	2715
Qy	837	LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla	856
Db	2716	TTGCCGACTGGGACTGGACACGGCTGCATCCGCTGTGAGAAAGATGGAGTTTATTTGCT	2775
Qy	857	ValLeuCystIleGlnThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer	876
Db	2776	GTGCTCTGCATAGAAACCGCCACTATGTGTGCTTTTGTGAAGTACGGGAAGGATGACTCT	2835
Qy	877	AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle	896
Db	2836	GCCTGGCTCTTCTTTGACAGCATGGCTGATCGAGATGGTGGTCAAGATGGCTTCAACATT	2895
Qy	897	ProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeu	916
Db	2896	CCACAAGTGACACCTGCCCAAGAGTAGGAGAGTACTTGAAGATGTCTCTCGAGGACCTG	2955

QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyr 936
DB 2956 CACTCTTTGGACTCCAGAGAGATTCAAGGCTGTGCGCGAGACTTCTTTGGCATGCATAC 3015

QY 937 MetCysMetTyrClnSerProThrMetSerLeuTyrIlys 949
DB 3016 ATGTGATGTACAGAGTCAACCATGAGCTGTATACAAA 3054

RESULT 12
BD160617 2845 bp DNA linear PAT 17-JAN-2003
LOCUS
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD160617
VERSION BD160617.1 GI:27866375
KEYWORDS JP 2002191363-A/15460.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2845)
Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,
Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 15460 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/15460
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT CDS Location/Qualifiers
1..2845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2845
Score: 4601.00 Matches: 875
Percent Similarity: 99.32% Conservative: 0
Best Local Similarity: 99.32% Mismatches: 2
Query Match: 91.40% Indels: 4
DB: 6 Gaps: 4

US-09-671-687A-3 (1-949) x BD160617 (1-2845)

QY 73 LysIleLeuGluGlnProHisAlaValLeuPheValAspGlu---AspValValGluIle 91
DB 2 AAAATTCAGAGCAACTCATGCAGTCTCTTTGTTGATGAAAAGGATGTTGTAGAGATA 61

QY 92 AsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGluGluArgPheSerLeu 111
DB 62 AATGAAAAGTTTCAGAGTTACTTTTGGCAATTACCAATTTGTGAGGAGAGTTTCAGCCTG 121

QY 112 PheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGlyCysProValIlys 131
DB 122 TTTAAAAACAGAAACAGACTAAGTAAGGGCTCTCAATAGACGTGGGTCTCTGTGAAA 181

QY 132 ValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPheArgGlyProLeu 151
DB 182 GTACAGCTGAGATCTGGGGAGAAAAAATTTCTCTGGAGTTGTACGCTTCAGAGGACCCCTG 241

QY 152 LeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeuGluGlyArg 171
DB 242 TTAGCAGAGAGACAGCTCTCCGGAATATTCTTTGGAGTTGAAATTGCTGGAAGAGGCTGT 301

QY 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAsp 191
DB 302 GGTCAAGGTTTCACTGACGGGTGTACCAAGGGGAAACAGCTTTTTCAGTGTGATGAAGAT 361

QY 192 CysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeu 210
DB 362 TGTGGCGTGTGTTGTCATTTGGCAAGCTAGAACTCATAGAGATGATGACACTGCATTG 421

QY 211 GluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProLeuGluIle 230
DB 422 GAAAGTATTACGACAGTCTCTGGGACACAATCAGGTGCAACTCTCTCTCTTTGGAAATA 481

QY 231 AsnSerArgValSerLeuLysGlyGluThrIleGluSerGlyThrValIlePheCys 250
DB 482 AACTCCAGAGTTTCTTTTGAAGGTGGAGAAACAATAAGATCTCGAACAGTTATATTTCTGT 541

QY 251 AspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnPro 270
DB 542 GATGTTTTGCCAGAAAGAAAGCTTAGGATATTTTGTGTTGGTGGACATGGATAACCCCT 601

QY 271 IleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGlu 289
DB 602 ATTGGCAACTGGATGGAGATTTTGTGAGTGCAGCTTTGTAGTTTGGTGTGTTGAA 661

QY 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArg 309
DB 662 AGTACAAATCTATTGTCACATCAATGATATCATCCAGAGAGTGTGACGAGGAAGGAGG 721

QY 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsn 329
DB 722 CTTCCCAAACTGGCTTTTATGTCAAGAGGTGTTGGGGACAAGGTTTCATCCAGTCAAT 781

QY 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPhe 348
DB 782 AAACCAAGGCTACAGAGTCTACTCAGACCTCGAAATAGAAACAGATCTGAATTTATTT 841

QY 349 TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrp 368
DB 842 TATACCTTAAATGGGTCTTCTGTGACTCACAACCAATCCAAATCAAATAATACATGG 901

QY 369 TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe 388
DB 902 TACATTGATGAAGTTGCAGAAAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTT 961

QY 389 AspArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArg 408
DB 962 GACCGTTTTCACACACACTCCAGCTCTCTCTGTGACTCCTGACACCGAGAACAGA 1021

QY 409 PheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHis 428
DB 1022 TTCCACTCTTTTACCATTCAGTCTCACAAGATGCCCAATACCAATGGAAGTATTGGCCAC 1081

QY 429 SerProLeuSerLeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGln 448
DB 1082 AGTCCACTTTCTCTGTGAGCCCACTGTCTCTGGGAATCATCATGTTTGAAGTGGGTCTATTG 1141

QY 449 GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeu 468
DB 1142 GAGAGTCCACCTTGGCCATGCTCTCTGGGAATCATCATGTTTGAAGTGGGTCTATTG 1201

QY 469 AlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProPro 488
DB 1202 GCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAACTCCGTTGGATCGGTGACCCACA 1261

QY 489 GlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAsp 508
DB 1262 GGACTGAATGAAGTGTCTGCTGGAGCTGGAACCTGGAAGATGAGTGTGAGGCTGTACCGAT 1321

Db 2 AAAATTCAGACCACTCATGCAGTCTCTTTGTTGATGAAAGGATGTTGTAGAGATA 61
Qy 92 AnGluLyPheThrGluLeuLeuAlaIleThrAsnCysGluGluArgPheSerLeu 111
Db 62 AATGAAAAGTTTACAGAGTTACTTTTGGCAATTAACCAATTTGTGAGGAGGTTTCAGCCTG 121
Qy 112 PheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGlyCysProValLys 131
Db 122 TTTTAAAAACAGAAACAGACTAAGTAAGGCTTCCAAATAGACGTGGGCTGTCTGTGAA 181
Qy 132 ValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPheArgGlyProLeu 151
Db 182 GTACAGCTGAGATCTGGGAGAGAAATTTCTGGAGTTGTACGCTTCAGAGACCCCTG 241
Qy 152 LeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeuGluGlyArg 171
Db 242 TTAGCAGAGAGACAGTCTCCGGAATATTTCTTTGGAGTTGAATTTCTGGAAGAAGTCTGT 301
Qy 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAsp 191
Db 302 GGTCAAGGTTTCACTGACGGGTGTACCAAGGGAAACAGCTTTTTCAGTGTGATGAAGAT 361
Qy 192 CysGly- --PheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeu 210
Db 362 TGTGGCGTGTGTTGTCATTTGGACAAGCTAGAACTCATGAAGATGATGACACTGCATTG 421
Qy 211 GluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProProLeuGluIle 230
Db 422 GAAAGTGATTCGCAGGTCTGGGACACATGACAGTGCNACTTCTCTTTGGAAATA 481
Qy 231 AsnSerArgValSerLeuLysGlyGluThrIleGluSerGlyThrValIlePheCys 250
Db 482 AACTCCAGAGTTTCTTTGAAGTTGGAGAAACAATAGAACTCTGGAACAGTTATATTCTGT 541
Qy 251 AspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnPro 270
Db 542 GATGTTTTCCAGGAAAGAAAGCTTAGGATATTTTGGTGTGGACATGGATAACCCCT 601
Qy 271 IleGlyAsnTrpAspGlyArgPheAspGlyVal- --LeuCysSerPheAlaCysValGlu 289
Db 602 ATTGGCACTGGAGTGAAGATTGATGAGTGACGCTTTGTAGTTTGGGTGTGTGAA 661
Qy 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArg 309
Db 662 AGTACAAATCTATTGCATCAATGATATCATCCAGAGAGTGTGACGAGGAAAGAGG 721
Qy 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsn 329
Db 722 CCTCCAAACTTTGCCCTTTATGTCAAGAGGTGTTGGGACAAAGGTTTCATCCAGTCAAT 781
Qy 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArg- --ArgSerGluLeuPhe 348
Db 782 AAACCAAGGCTACAGATCTACTCAGACCTCGGAAATAGAAACAGATCTGAATATTAT 841
Qy 349 TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrp 368
Db 842 TATACCTTAATGGGTCTTCTGTGACTCACAACCACTCCAAATCAAAATACATGG 901
Qy 369 TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe 388
Db 902 TACATTGATGAGTTGCAGAAAGCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTT 961
Qy 389 AspArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArg 408
Db 962 GACCGTCTTCCACCACTCCAGCTCTCCCTGTTGAACCTCAGACCCCGAAGACAGA 1021
Qy 409 PheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHis 428
Db 1022 TTCCACTCTTACCATTTAGTCTCACCAGATGCCCCCAATACCAATGGAGTATTTGCCAC 1081
Qy 429 SerProLeuSerLeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGln 448
Db 1082 AGTCCACTTTTCTGTGAGCCAGTCTGTAATGGAAGAGCTAAACACTGCAACCCGTCCTCAA 1141

Qy 449 GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeu 468
Db 1142 GAGAGTCCACCTTGGCCATGCTCTCTGGAACTCACAATGCTAGAAAGTGGGCTCATTTG 1201
Qy 469 AlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProPro 488
Db 1202 GCTGAAGTTTAAAGAGAACCTTCTCTTATGGGGTAATCCGTTGGATCGGTGACGCCACCA 1261
Qy 489 GlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAsp 508
Db 1262 GGAAGTGAATGAGTGTCTCGTGGACTGGAACTGGAAAGATGAGTGTGCAGGCTGTACCGAT 1321
Qy 509 GlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLys 528
Db 1322 GGAACCTTCAGAGGCACTCGGTATTTTACCTGTGCCCTGAAGAAGCGCTGTTTGTGAAA 1381
Qy 529 LeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIle 548
Db 1382 CTGAAGAGCTGACGGCTGACTCTAGGTTTGCATCATTTGACGCGGTTTCCAAATCAGATT 1441
Qy 549 GluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThr 568
Db 1442 GAGCGCTGTAACTCTTTAGCAITTTGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACT 1501
Qy 569 ProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGln 588
Db 1502 CCACCAAAATGGAAAAGAGGCTTGGAGATAATGATTTGGGAAGAAAGGATCCAG 1561
Qy 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
Db 1562 GGTCAATCAAAATTTGTTACTTTAGACTCAACCTTATTTCTGCTTATTTTGTGTTTCT 1621
Qy 609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
Db 1622 GTTCTGGACACTGTGTACTTGGACCCCAAGAAAGAACGATGTAGAATATATTAGTGAA 1681
Qy 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
Db 1682 ACCCAAGACTACTGAGGACAGAAATTTGTAATCTCTGAGATAATATGATATGTGTGT 1741
Qy 649 AlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaLaserGlyPhe 668
Db 1742 GCCACAAAATTTATGAACTGAGGAAAATACTTTGAAAAGGTGGAGGCTGCATCAGGATT 1801
Qy 669 ThrSerGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg 688
Db 1802 ACCTCTGAAGAAAAGATCCTGAGGAATTTCTGAATATTTCTGTTTCATCATATTTTAAAG 1861
Qy 689 ValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyr 708
Db 1862 GTAGAACCCTTTGTTAAATTAAGATCAGCAGGTCAAAAGGTACAGATTTGTACTTCTAT 1921
Qy 709 GlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGlu 728
Db 1922 CAAATTTTATGAAAAAATAGAAAGTTGGCGTTCCCAAAATTCAGCAGTTGTTAGAA 1981
Qy 729 TrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGln 748
Db 1982 TGGTCTTTTATCAACAGTAACTCGAAATTTGACAGAGCACCATCATGCTCTGATTATTTCAG 2041
Qy 749 MetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeu 768
Db 2042 ATGCTTCGATTGGAAAAGACTTTAACTATTTTAAAAAATTTTCTTCTCTGGAATTA 2101
Qy 769 AsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAla 788
Db 2102 AATATAACAGATTTACTTGAAGACACTCCACAGACAGTGGCGGATATGTCGAGGCGCTTGA 2161
Qy 789 MetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGln 808
Db 2162 ATGTATGAGTGTAGAAATGCTACGAGATCCGACATCTCAGCTGGAATAATCAAGCAG 2221

QY 809 PheCysLysThrCysAsnThrGlnValHisLeuHisProIysArgLeuAenHisLysTyr 828
 Db 2222 TTTTGTAAACCTGCAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAAT 2281
 QY 829 AsnProValSerLeuProLysAspLeuProAspTTPAspTTPArgHisGlyCysIlePro 848
 Db 2282 AACCCAGTGTCACTTCCCAAGACTTACCGACTGGGACTGGAGACACGGCTGCATCCCT 2341
 QY 849 CysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPhe 868
 Db 2342 TGCCAGAAATAGAGTATTGTGCTCTCTGTCATGAAACCAAGCCACTATGTGTGCTTT 2401
 QY 869 ValLysTyrGlyIysAspSerAlaTTPLeuPhePheAspSerMetAlaAspArgAsp 888
 Db 2402 GTGAAGTATGGGAAGACCATTCCTGGCTCTCTTTTGACAGCATGGCCGATCGGGAT 2461
 QY 889 GlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyr 908
 Db 2462 GTGTGTCAGAAAGTCTTCAACATTCCTCAAGTCAACCCATGCCCAGAGTAGAGAGATAC 2521
 QY 909 LeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAla 928
 Db 2522 TTGAAGATGCTCTGGAAGACCTGCAATTCCTTGACTCCAGGAGAAATCAAGGCTGTGCA 2581
 QY 929 ArgArgLeuLeuCysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyr 948
 Db 2582 CGAAGACTCTTTGTGATCATATATGTGATGTACAGAGTCCAAACAAATGAGTTGTGATC 2641
 QY 949 Lys 949
 Db 2642 AAA 2644

RESULT 14
 AK024348
 LOCUS Homo sapiens cDNA FL114286 fis, clone PLACE1006129, highly similar
 DEFINITION to Homo sapiens HSPC057 mRNA.
 ACCESSION AK024348
 VERSION 1
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Negai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirakawa, S., Chiba, Y., Iehida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Coto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shiget, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,

Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
 Complete sequencing and characterization of 21,243 full-length human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004) - late
 14702039
 2
 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Iehida, S., Murakami, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Iehii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 2845)
 Isogai, T. and Otsuki, T.
 Direct Submission
 Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
 FEATURES
 Location/Qualifiers
 1..2845
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1006129"
 /tissue_type="placenta"
 /clone_lib="PLACE1"
 /notes="cloning vector: pME18SFL3"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2845
 Score: 4601.00 Matches: 875
 Percent Similarity: 99.32% Conservatives: 0
 Best Local Similarity: 99.32% Mismatches: 2
 Query Match: 91.40% Indels: 4
 DB: 9 Gaps: 4
 US-09-671-687A-3 (1-949) x AK024348 (1-2845)
 QY 73 LysIleLeuGluGlnProHisAlaValLeuPheValAspGlu---AspValValGluIle 91
 Db 2 AAATTTCTAGAGCAACCTCATGAGTCTCTTTTGTGATGAAGAAGATGTTGTAGAGATA 61
 QY 92 AsnGluLysPheThrGluLeuLeuLeuAlaIleThrAsnCysGluGluArgPheSerLeu 111
 Db 62 AATGAAAAGTTTACAGAGTTTACTTTTGGCAATTACCAATTGTGTGAGGAGAGTTTCAGCCTG 121
 QY 112 PheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGlyCysProValLys 131
 Db 122 TTTTAAAAACAGAAACAGACTTAAGTAAAGGCCTCCAAATAGACGCTGGCTGTCTGTGAAA 181
 QY 132 ValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPheArgGlyProLeu 151
 Db 182 GTACAGCTGAGATCTGGGGAAGAAAATAATTTCTTGGAGTTGTACGCTTCAGAGGACCCCTG 241
 QY 152 LeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeuGluGlyArg 171
 Db 242 TTAGCAGAGAGGACAGCTCTCCGGAATATTCTTTGGAGTTGAAATTTGCTGGAAGAAGTTCGT 301
 QY 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAsp 191
 Db 302 GGTCAAGGTTTCACTACGCGGGGTGTACCAAGGGGAAACAGACTTTTTCAGTGTGTGATGAAGAT 361

Qy	192	CysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeu	210	Db	1442	GAGCGCTGTAACCTCTTTAGCATTTGGAGCTACTTAAGTGAAGTAGTAGAAGAAATACT	1501
Db	362	TGTGGCGTGTGTGTGCAATGGACAAGCTTAGAACTCATAGAAGATGATGACACTGCATTG	421	Qy	569	ProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGln	588
Qy	211	GluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProProLeuGluIle	230	Db	1502	CCACCACAAANTGGAAAAGAGCGCTTGGAGATAATGATTTGGGAAGAAAGAGCATCCAG	1561
Db	422	GAAAGTGATTAACGCAAGGCTCGGGACACAATGCAAGGTCGAACTTCCTCTTTGGAAATA	481	Qy	589	GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer	608
Qy	231	AsnSerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCys	250	Db	1562	GGTCATTACAATCTCTTTACTTTAGACTCAACCTTAATCTGCTTAATTTGCTTTTGTCT	1621
Db	482	AACTCCAGAGTTCTTTGAGGTTGGAGAACAAATAGAACTCTGGAAACGTTATATCTGT	541	Qy	609	ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu	628
Qy	251	AspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAspPro	270	Db	1622	GTTCTGGACACTGTGTACTTGGACCCAAAGAAAAGACGATAGAAATATTAGTGAA	1681
Db	542	GATGTTTTGCCAGGAAAAGAAAGCTTAGGATATTTGTTGGTGTGGACATGATACCCCT	601	Qy	629	ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys	648
Qy	271	IleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGlu	289	Db	1682	ACCAAGAGCTACTGAGGACAGAAATTTGTAATCTCTGAGAAATATATGGATATGTGT	1741
Db	602	ATTGGCAACTGGATGGAGATTTGATGGAGTGCAGCTTTGTAGTTTTCGCTGTGTGAA	661	Qy	649	AlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAspGlyPhe	668
Qy	290	SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArg	309	Db	1742	GCCACAAAATTTATGAACTGAGGAAAATACITGAAAAGGTGGAGCTGCATCAGGATTT	1801
Db	662	AGTACAAATCTATTGCACATCAATGATATCATCCACAGAGTGTGACGAGAAAGAGG	721	Qy	669	ThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg	688
Qy	310	ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsn	329	Db	1802	ACCTCTGAAAGAAAAGATCCTGAGGAATCTTGAATATTTCTGTTTCATCATATTTTAAG	1861
Db	722	CCTCCCAACTTGCCTTTATGTCAAGAGGTGTGGGACAAAAGTTTCATCCAGTCATAAT	781	Qy	689	ValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyr	708
Qy	330	LysProLysAlaThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPhe	348	Db	1862	GTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGTACAAGATTGTACTTCTAT	1921
Db	782	AAACCAAGGCTACAGGATCTACCTCAGACCCTGGAAATAGAAACAGATCTGAAATATTTT	841	Qy	709	GlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGlu	728
Qy	349	TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrp	368	Db	1922	CAAAATTTTATGGAAAAAATGGAAGTTGGGTTCCACAAATTCAGCAGTTGTTAGAA	1981
Db	842	TATACCTTAATAGGGTCTCTGTGTACTCACACCAACCAATCCAAATCCAAATAATACATG	901	Qy	729	TrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGln	748
Qy	369	TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe	388	Db	1982	TGCTCTTTTATCAACAGTAACCTGAAATTTGCAGAGCACCACATCTCTGATTTATTCAG	2041
Db	902	TACATTTGATCAAGTTGCAGAGACCTCGCAAAATCTCTTACAGATATCTACAGACTTT	961	Qy	749	MetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuLeuLeu	768
Qy	389	AspArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArg	408	Db	2042	ATGCTCTCGAATTTGGAAAAGACTTTAAACTATTTTAAAAAAATTTTCTCTCTGGAAATTA	2101
Db	962	GACCGTCTTCCACCACCTCCAGCCTCTCTCTGTGAACTCACTGACCCACCGAGAACAGA	1021	Qy	769	AsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAla	788
Qy	409	PheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHis	428	Db	2102	ATAATAACAGATTTACTTGAAGACACTCCACAGACAGTCCGGATATGTGGAGGCTTGCA	2161
Db	1022	TTCCACTCTTTACCAATTCAGTCTCACCAAGATGCCCAATACCAATGGAAGTATTTGCCAC	1081	Qy	789	MetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGln	808
Qy	429	SerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGln	448	Db	2162	ATGTATGATGTAGAGATGCTACGACGATCCGACATCTCAGCTGGAATAAATCAAGCAG	2221
Db	1082	AGTCCCACTTTCTGTGAGCCCACTGTGTAATGGAAGAGCTAAACACTGCACCCGTCCAA	1141	Qy	809	PheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyr	828
Qy	449	GluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeu	468	Db	2222	TTTTGTAAACCTTCGAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATATAATAT	2281
Db	1142	GAGAGTCCACCTTGGCCATGCTCTCTGGAACTCACTGCTAGAAAGTGAGCTCATTTG	1201	Qy	829	AsnProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIlePro	848
Qy	469	AlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlnProPro	488	Db	2282	AACTCCAGTGTCACTTCCCAAGACTTTACCCGACTGGGACTGGAGACACGGCTGCATCCCT	2341
Db	1202	GCTGAAGTTAAGGAGAACCTCTCTTTCTATGGGGTAAATCCGTTGGATCGGTACGCCCA	1261	Qy	849	CysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPhe	868
Qy	489	GlyLeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAsp	508	Db	2342	TGCCAGATATGAGTATTTGCTGTTCTCTGCATAGAAAACCAAGCCACTATGTTGCTTTT	2401
Db	1262	GGACTGAATGAAGTGTCTGCTGGACTGGAACTGGAAAGATGAGTGTGCGAGCTGTACGGAT	1321	Qy	869	ValLysTyrGlyLysAspAspSerAlaTrpIlePhePheAspSerMetAlaAspArgAsp	888
Qy	509	GlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLys	528	Db	2402	GTCAAGTATGGGAAGGACGATTTCTGCTGGCTCTTCTTTCAGACATGGCCGATCGGGAT	2461
Db	1322	GGAACTTCAGAGGCACTCGTATTTTCACTGTGCCCTGGAAGGGCGCTGTTGTGAA	1381	Qy	889	GlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyr	908
Qy	529	LeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIle	548	Db	2462	GGTGGTTCAGAAATGGCTTCAACATTTCTCAAGTACCCTCATCCCAAGAAAGTAGAGATAC	2521
Db	1382	CTGAAGAGCTGCAGGCGTACTAGGTTTGATCATTTGACGCGGTTTCCAAATCAGATT	1441	Qy	909	LeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAla	928
Qy	549	GluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThr	568				

Db 2522 TTGAAGATGCTCTCTGGAAGACCTGCGATTCTTGGACTCCAGGAGATCCAAAGCTGTGCA 2581

QY 929 ArgA9LeuLeuCuAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyr 948
|||||

Db 2582 CGAAGACTCTTTGTGATGCATATATGTGTGCTACGAGTCCAAATGAGTTGTATC 2641

QY 949 Lys 949
|||

Db 2642 AAA 2644

RESULT 15
AC123449

LOCUS Rattus norvegicus clone CH230-21D6, *** SEQUENCING IN PROGRESS ***,
DEFINITION 3 unordered pieces.

AC123449

ACCESSION AC123449.3 GI:23265406

VERSION HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 241990)

AUTHORS Muzny,D,Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,
Hollins,B., Howells,S., Huiy,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Mirza,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quirroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P.,
Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,J., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs, R.A.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 241990)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 241990)

AUTHORS Rat Genome Sequencing Consortium.

TITLE Direct Submission

JOURNAL Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT On Sep 21, 2002 this sequence version replaced gi:21902787.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GXOF

Center clone name: CH230-21D6

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 229402 bases at least Q40

Consensus quality: 231763 bases at least Q30

Consensus quality: 233057 bases at least Q20

Estimated insert size: 260790; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

* NOTE: This sequence may represent more than one 'clone'.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 238985: contig of 238985 bp in length
* 238986 239085: gap of unknown length
* 239086 240731: contig of 1646 bp in length
* 240732 240831: gap of unknown length
* 240832 241990: contig of 1159 bp in length.

FEATURES

source

1. .241990
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-21D6"
1323. .2205
/notes="clone boundary
clone end:Sp6
site:EcORI
end_sequence:BH269859"
complement(236297. .237124)
/notes="clone boundary
clone end:T7
site:EcORI
end_sequence:BH269858"

misc_feature

misc_feature

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	241990
Score:	4600.50	Matches:	883
Percent Similarity:	95.29%	Conservative:	27
Best Local Similarity:	92.46%	Mismatches:	38
Query Match:	91.39%	Indels:	8
DB:	2	Gaps:	5
US-09-671-687A-3 (1-949) x AC123449 (1-241990)			
QY	1	MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTyrGluGluArgIle	20
DB	205375	ATGAGTTTACGGCTGTGGAAACCAAGAGAAGTTACTTCCCTTCTGGGAAGAACGGCTT	205434
QY	21	PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
DB	205435	TTTTATCTCTCTTCAAGAATGCAGTGAAC---AAACAGACACAGAAAGCTCTCTGAGA	205490
QY	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisArgIle	60
DB	205491	GTACCCAAAGGAGCATAGACAGTACATCCAAAGACTGTTCCGTGGGCAATTCACAGATT	205550
QY	61	ProSerAlaLysGlyLysLysGlnIleGlyLeuLysIleLeuGlnGlnProHisAla	80
DB	205551	CCTTCTGCTTAAAGGCAAGAAATAATCAGATTGGATTAAATAATCTTAGCAACACCGCATCA	205610
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99
DB	205611	GTTCTGTTTGTGATGAAGAAGNTGTTAGAAATAAATAAATAATTCACAGATTACTG	205670
QY	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
DB	205671	TTGGCAATTTACCAACTGTGAGGAGAGGCTCAGCCTATTTAGAAACAGAAATCCGACTAAGT	205730
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
DB	205731	AAAGGCTCCAGGTAGACGTGGGCAGTCCCTGTGAGAGTACAGCTGGCATCTGGGGAGGAG	205790
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
DB	205791	AGATTTCAGGAGTTGTAGCTTTACAGGAGACCTTTATTACGGAGAGACGGTGTCCGGG	205850
QY	160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal	179
DB	205851	ATTTCTTTGGAGTAGAATTTACTGGAAAGGTCGTGGCCAAAGTTTCACTGATGGGTG	205910
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
DB	205911	TATCAAGGAAACAGCTCTTCCAGTGTGATGAGGACTGTGGTGTGTTGTTGCAATTGGAC	205970
QY	199	LysLeuLeuLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218
DB	205971	AAGCTGGAGCTTATAGAAGATGATGACAAATGGGTTGGAAGTGAATTTTGCAGGCCACGGA	206030
QY	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238
DB	206031	GATACAGTCCAGGTTGAACCTCCCTTTGGAAATAAATCTCCAGAGTTTCTTTGAAGGTT	206090
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258
DB	206091	GGAGAAAGTACAGAACTCTGAAACAGTGATATTCTGTGATGTTTTTACCAGGAAAGAGAT	206150
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278
DB	206151	CTAGGATATTGTTGTGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAAGGTTT	206210
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297
DB	206211	GATGGGCTTCAGGCTTTCAGTTTGCAGTGTGAGAGTACAGTTCTCTTACACATCAAT	206270
QY	298	AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer	317
DB	206271	GACATCATCCAGATAGCGTGACACAGGAAAGGAGACCTCCCAAACTTGCCTTTATGTCA	206330
QY	318	ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr	337
DB	206331	AGAGGTGTAGTGCACAAAGGCTCATCTAGTCAATAAACAAGGTTTACAGGATCTACC	206390
QY	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356
DB	206391	TCAGACCTTGGAGTAGAACAAGATCTGAATATTTTATACCTTAAATGGGTCTATCTGTT	206450
QY	357	AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp	376
DB	206451	GACTCACACAACAATCCAAGTCCAAAACCCATGTTACATTGATGAAGTTGCAGAAGAC	206510
QY	377	ProAlaLysSerLeuThrGluLysSerThrAspPheAspArgSerSerProProLeuGln	396
DB	206511	CCTGCAAGTCACTTACAGAGATGTTCTCAGACTTCGGACATTCATCGCCTCCACCCCAA	206570
QY	397	ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu	416
DB	206571	CCTCCTTCATGAACCTCTTGTCTAGCGAGAACAGATTCCACTCCTTACCTTCAGCCTG	206630
QY	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerPro---LeuSerLeuSerAlaG	436
DB	206631	ACAAAGATGCCCAACACTAATGTCAGCAGTGGCTCAGAGTCCACTCTCTCTGTGAGTGC	206690
QY	436	InSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetP	456
DB	206691	AGTCTGTGATGGGGAGCTGAACAGCAGCGCTGTCCAGGAGAGTCCACCTTGGCCAGCT	206750
QY	456	roProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProp	476
DB	206751	CTTCTGGGAATGGACACGGGCTAGAGGTGGGCTCACTGGCTGAAGTAAAGAGAACCCCC	206810
QY	476	roPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLeuAlaG	496
DB	206811	CGTTCTATGGGGTATTCGGTTGGATTGGCCAGCCAGGAGGCTCAGAGCGTCTCGCTG	206870
QY	496	lyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgT	516
DB	206871	GATTGGAAACCGGAAGATGAATGCGCAGGTTGCAGCGATGGAACCTTTCAGGGCACCGCT	206930
QY	516	yrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspS	536
DB	206931	ATTTCACCTGTGCTTGAAGAAAGCACCTGTTCTGTGAACCTGAAAGAGTGCAGACAGACT	206990
QY	536	erArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaP	556
DB	206991	CTAGGTTTGCATCTTGCAGCCTGTTTCCAATCAGATCGAAGAGGTGAATCTTTTAGCAT	207050
QY	556	heGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluG	576
DB	207051	TTGGGGGCTACTTAAAGTAGTAGTAGAAGAAATAACGCCACCTTAAATGGAAAAGGAAG	207110
QY	576	lyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrL	596
DB	207111	GTTTAGAGATAATGATTGGAAAGAAAGAGCATCCAGGGCCATTACAAATCTTTGTTACT	207170
QY	596	euAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuA	616
DB	207171	TAGACTCAACTTTATTTCTGCTTATTTGCTTTTGTCTGCTGCGCTGGACACTGTATTTACTTA	207230
QY	616	rgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrG	636
DB	207231	GACCCAAAGAGAGAAATGAGTAGAGTATTTACAGTGAGACTCAAGAGCTACTGAGGACAG	207290
QY	636	luIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuA	656
DB	207291	AGATAGTCAATCTCTGAGAAATATATGATATGATGTGTGTGTCGACCAAGAGATTATGAAGCTGA	207350
QY	656	rgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProG	676
DB	207351	GGAAAAATCTTGAAGAAAGTTGAGGCTCATCAGGATTTTACCTCTGAGGAAAAAGATCTCTG	207410
QY	676	luGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleA	696


```
Db 207411 |AAGATTTCTAAACATCCTGTTTCATGATATTTTAAAGGATGAACCATTTGTTAAAAATAT 207470
QY 696 |rSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG 716
Db 207471 |GGTCACAGGTCAAAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGGAATAAATG 207530
QY 716 |lulysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnL 736
Db 207531 |AGAAAGTCGGAGTACCACAAATCCACAGCTATTAGAAATGTTCTTTTATCAACAGCAACC 207590
QY 736 |eulysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspP 756
Db 207591 |TGAAATTTCCGGAGGACCATCATGTGTTGATATCCAGATGCCCTTGGTTTGGAAAGACT 207650
QY 756 |helysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluA 776
Db 207651 |TTAAACTATTAAAAAATTTTTCCTTCCTCGGAATTAATATACAGATTACTTTGAAG 207710
QY 776 |spThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysT 796
Db 207711 |ACACTCCAGGCGAGTCGCGCATCTGTGGAGACTCGCCATGTATGAGTGTAGAGAGTGCT 207770
QY 796 |yrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG 816
Db 207771 |ATGATGCCCGGACATCTCGCAGGGAAGATCAAGCAGTTCCTGAAGACCTGCAGCACTC 207830
QY 816 |lnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysA 836
Db 207831 |AGGTTACCTTCATCCAGAGACTGAATCACACTTACCATCCAGTATCACATCCCAAG 207890
QY 836 |spLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheA 856
Db 207891 |ACTTGCCCGACTGGGACTGGAGACATGGCTGCATCCCATGTCAGAAAGATGGAGTTATTG 207950
QY 856 |lavallLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAs 876
Db 207951 |CTGTGCTCTGCATAGAAACCAGCCACTATGTTGCTTTTGTGAAGTACGGGAAGGATGACT 208010
QY 876 |erAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI 896
Db 208011 |CTGCCTGGCTCTCTTTTGACAGCATGGCTGATCGAGATGGTGTGAGATGGCTTCAACA 208070
QY 896 |leProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspL 916
Db 208071 |TTCCACAAGTGACACCCTGCCCAAGAGTAGGAGAGTACTTTGAAGATGTTCTCTGGAGACC 208130
QY 916 |euHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaT 936
Db 208131 |TGCACCTCTTTGGACTCCAGAAGGATTCAAGGCTGTGGCGCAGACTTCTTTGGGATGCAT 208190
QY 936 |yrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 208191 |ACATGTGCATGTACCAAGTCCAAACCATGAGCTGTGACAAA 208231
```

Search completed: April 18, 2005, 18:28:45
Job time : 9311 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:23:45 ; Search time 54 Seconds

(without alignments)
1242.248 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSFYWEERI.....RLLCDAYMCYQSPWLSLYK 949

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 308579 seqs, 70686408 residues

Total number of hits satisfying chosen parameters: 308579

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pending Patents AA New.*

2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*

3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*

4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*

5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*

6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*

7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep.*

8: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	4971.5	98.8	956	US-10-760-678-4	Sequence 4, Appl
2	122.5	2.4	1499	1 PCT-US04-24868-15	Sequence 15, Appl
3	122	2.4	2296	6 US-10-696-909A-46	Sequence 46, Appl
4	122	2.4	2752	6 US-10-696-909A-44	Sequence 44, Appl
5	120.5	2.4	915	8 US-60-655-875-141720	Sequence 141720
6	119.5	2.4	1451	1 PCT-US05-03344-1	Sequence 1, Appl
7	119.5	2.4	1451	7 US-11-046-346-1	Sequence 1, Appl
8	119.5	2.4	2412	8 US-60-643-717-7896	Sequence 7896, Ap
9	115	2.3	2543	6 US-10-450-763-51654	Sequence 51654, A
10	113	2.2	1181	8 US-60-664-582-558	Sequence 558, App
11	113	2.2	1195	8 US-60-659-397-1215	Sequence 1215, Ap
12	113	2.2	1200	8 US-60-664-582-563	Sequence 563, App
13	113	2.2	1213	7 US-11-033-545-572	Sequence 572, App
14	113	2.2	1213	8 US-60-659-397-1214	Sequence 1214, Ap
15	113	2.2	1226	7 US-11-033-545-313	Sequence 313, App
16	113	2.2	1226	8 US-60-664-582-562	Sequence 562, App
17	113	2.2	1226	8 US-60-664-582-564	Sequence 564, App
18	112.5	2.2	1116	6 US-10-818-066-31	Sequence 31, Appl
19	112	2.2	807	6 US-10-450-763-45311	Sequence 45311, A
20	112	2.2	869	7 US-11-031-175-10125	Sequence 10125, A
21	112	2.2	3684	1 PCT-US05-01768-8	Sequence 8, Appl
22	111.5	2.2	1205	6 US-10-489-448-1764	Sequence 1764, Ap
23	111.5	2.2	1304	8 US-60-651-509-263	Sequence 263, App
24	111.5	2.2	1304	8 US-60-651-235-1508	Sequence 1508, Ap
25	111.5	2.2	1304	8 US-60-664-579-738	Sequence 738, App

26	111	2.2	1673	8	US-60-643-717-14226	Sequence 14226, A
27	110	2.2	1497	6	US-10-450-763-51156	Sequence 51156, A
28	110	2.2	2644	6	US-10-450-763-50462	Sequence 50462, A
29	109.5	2.2	1304	8	US-60-651-509-262	Sequence 262, App
30	109.5	2.2	1304	8	US-60-651-235-1507	Sequence 1507, Ap
31	109.5	2.2	1304	8	US-60-664-579-737	Sequence 737, App
32	109	2.2	452	1	PCT-US04-43356-148	Sequence 148, App
33	109	2.2	452	7	US-11-021-949-148	Sequence 148, App
34	109	2.2	452	7	US-11-021-328A-45	Sequence 45, Appl
35	109	2.2	521	7	US-11-044-879-4	Sequence 4, Appl
36	109	2.2	866	1	PCT-US04-17965-1594	Sequence 1594, Ap
37	109	2.2	866	1	PCT-US04-17965B-1594	Sequence 1594, Ap
38	109	2.2	866	1	PCT-US04-17965C-1594	Sequence 1594, Ap
39	109	2.2	866	6	US-10-863-905-1594	Sequence 1594, Ap
40	108.5	2.2	493	8	US-60-651-509-442	Sequence 442, App
41	108.5	2.2	493	8	US-60-651-235-1229	Sequence 1229, Ap
42	108.5	2.2	703	7	US-11-096-039-11	Sequence 11, Appl
43	108.5	2.2	809	6	US-10-450-763-40966	Sequence 40966, A
44	108.5	2.2	1205	6	US-10-450-763-52696	Sequence 52696, A
45	108	2.1	714	6	US-10-450-763-45506	Sequence 45506, A

ALIGNMENTS

RESULT 1

US-10-760-678-4

; Sequence 4, Application US/10760678

; GENERAL INFORMATION:

; APPLICANT: Derr, Jonathan

; APPLICANT: Fanslow, William

; APPLICANT: Dougall, William

; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING

; FILE REFERENCE: 3198

; CURRENT APPLICATION NUMBER: US/10/760, 678

; CURRENT FILING DATE: 2004-01-20

; PRIOR APPLICATION NUMBER: US/09/851, 673

; PRIOR FILING DATE: 2001-05-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 956

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-760-678-4

Query Match 98.8%; Score 4971.5; DB 6; Length 956;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY	1	MSSGLWSQEKVTSFYWEERIFYLLLOECSTVDKQTKLLKVPKSGISQYIQDRSVGHSRI	60
DB	1	MSSGLWSQEKVTSFYWEERIFYLLLOECSTVDKQTKLLKVPKSGISQYIQDRSVGHSRI	60
QY	61	PSAKGKNOIGLKILQPHAVLPVDE-DVVEINEKFTELLAIATNCERFSLFKNRRLS	119
DB	61	PSAKGKNOIGLKILQPHAVLPVDEKVVVEINEKFTELLAIATNCERFSLFKNRRLS	120
QY	120	KGLQIDVGCVPKVLRSGBEKGPGVVRFRGPLAAERTVSGIFPGVELLESGRGQGTGTV	179
DB	121	KGLQIDVGCVPKVLRSGBEKGPGVVRFRGPLAAERTVSGIFPGVELLESGRGQGTGTV	180
QY	180	YQCKQLFQCDQDCG-FVALDKLELIEDDDTALSDYAGPGDTMOWELPPLEINRSVLKG	238
DB	181	YQCKQLFQCDQDCG-FVALDKLELIEDDDTALSDYAGPGDTMOWELPPLEINRSVLKV	240
QY	239	GETIESGTVIFCDVLPKESLGYPGVGDMNDPIGNWDGDFGV-LCSFACVESTILLHIN	297
DB	241	GETIESGTVIFCDVLPKESLGYPGVGDMNDPIGNWDGDFGV-LCSFACVESTILLHIN	300
QY	298	DIIP-----ESVTQRRPPKLAFMRSRGVDKGSSHNKPKATGSTDSPGNR-RSELYTLNG	353
DB	301	DIIPALSESVTQRRPPKLAFMRSRGVDKGSSHNKPKATGSTDSPGNRNRSELFYTLNG	360

QY 302 ESWTQERRPPKLAFLMRGVDKSSS-----HNKPKATGTSF-----PGNRSELFTYTLNG 353
Db 1163 ESKEKALPPQEDATASPPRQDKFPFPVQDRPESLSLVFKDTRLTPPRERSGA-----G 1217
QY 354 SSVDSOPQ-----SKSKNTWYIDEVAEDPA-----KSLTEISTDFDRSSPPLQPP 398
Db 1218 SSPETKEQNSALPTSQDELMVEVEKSEBPAGQILSHLSSELKEMSTSNFESSPEVEER 1277
QY 399 PVNSLTENRHFSLPFLTKMPNTNGSIGHSPLSLSAQSVMEBLNAPVQES---PPLAM 455
Db 1278 PAVSLTLDSQSQASLEAVEVPSMASSWGGPHFS-----PEHKELNSPLRENSFGSPLEF 1333
QY 456 PPGNSHGLEV--GSLAEVKE--NPPF 477
Db 1334 RNSGPIGTMTNTGFSSEVKEDLNGPF 1359

RESULT 4

US-10-696-909A-44
; Sequence 44, Application US/10696909A
; GENERAL INFORMATION:
; APPLICANT: Lorens, James B.
; APPLICANT: Atchison, Robert E.
; APPLICANT: Frieria, Anabella
; APPLICANT: Holland, Sacha
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modulators of Angiogenesis and Tumorigenesis
; FILE REFERENCE: 021044-005820US
; CURRENT APPLICATION NUMBER: US/10/696,909A
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/512,251
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: US 60/421,989
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 2752
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: splicing coactivator subunit SRM300; RNA binding
; OTHER INFORMATION: protein; AT-rich element binding factor
US-10-696-909A-44

Query Match 2.4%; Score 122; DB 6; Length 2752;
Best Local Similarity 22.8%; Pred. No. 7.8;
Matches 88; Conservative 53; Mismatches 167; Indels 78; Gaps 16;
QY 128 CPVKVQLRSGEKEKFGVWRFRGPLLAERTVSGIFFGVVLELLE-BGRGQGFTDGVYQKQLF 186
Db 1016 CPOEKSQSLVQSCPSGLSICAGKVSSTPPGESYFGVSSQLKGQSTSPDH----- 1067
QY 187 QCDEDCGFVALDKLELIEDDTALEDYAGPGDTMQ--VELPPELINSRVSLKGGTIES 244
Db 1068 -----RSDTSSP-EVRQSHSESPSLQSKQSTSPKGRSRSS 1102
QY 245 GTVI-PCDVLPGKESIGYFVGVDN-DNPTGNWDGRPDGVLCSFACVESTILLHINDI-IP 301
Db 1103 SPVTELASRPIQRDRGEFSASPMKLSGMSPEQSRFQSDSSSYPTVDSNLLQGSRLTA 1162
QY 302 ESWTQERRPPKLAFLMRGVDKSSS-----HNKPKATGTSF-----PGNRSELFTYTLNG 353
Db 1163 ESKEKALPPQEDATASPPRQDKFPFPVQDRPESLSLVFKDTRLTPPRERSGA-----G 1217
QY 354 SSVDSOPQ-----SKSKNTWYIDEVAEDPA-----KSLTEISTDFDRSSPPLQPP 398
Db 1218 SSPETKEQNSALPTSQDELMVEVEKSEBPAGQILSHLSSELKEMSTSNFESSPEVEER 1277
QY 399 PVNSLTENRHFSLPFLTKMPNTNGSIGHSPLSLSAQSVMEBLNAPVQES---PPLAM 455
Db 1278 PAVSLTLDSQSQASLEAVEVPSMASSWGGPHFS-----PEHKELNSPLRENSFGSPLEF 1333

QY 456 PPGNSHGLEV--GSLAEVKE--NPPF 477
Db 1334 RNSGPIGTMTNTGFSSEVKEDLNGPF 1359
RESULT 5
US-60-655-875-141720
; Sequence 141720, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 141720
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna-SeqID 68143; Strand=-; Positions=1
; OTHER INFORMATION: -348,394-905,955-1966,2010-2170,2218-2308,2353-2976
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_001004509.1; Match level="QueryCover
; OTHER INFORMATION: 87%; HitCoverage=55%; E-value=1e-111; Identity=32%; Hit description
; OTHER INFORMATION: =si:buseml-18005.3 [Danio rerio] emb|CAD60788.1| novel protein sim
; OTHER INFORMATION: to DNA polymerases [Danio rerio]
US-60-655-875-141720

Query Match 2.4%; Score 120.5; DB 8; Length 915;
Best Local Similarity 21.9%; Pred. No. 2;
Matches 106; Conservative 69; Mismatches 195; Indels 115; Gaps 23;
QY 498 ELEDEECAGC--TDGTFRGTRYFTCALKKALFKLSCRPDSRFASLQVSNQIE----- 549
Db 49 ELSTDCEICCTGTGTFERTKEWLSKLSKFIFYQNSNRSEAD-GDEQPVAAVFEWILRAW 107
QY 550 -----RCNSLAFGY--LSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDS 598
Db 108 NNQYKTVIAHNASRFDGHEALNYICKTVRRPDVVMNGLKIYEFVRVHSPKSHMLIMRDS 167
QY 599 TLCLFAFSSVLTVLRLPKKNDVVEYSETQE--LLRTEIVNPLRIY--GYVCATKIMK 654
Db 168 CLIMPIPLADLKTTFNLDCEDKFPFPYAFNRRENYCIRLEKLPDQKMYEPMGSMKADKYEK 227
QY 655 LRKILEKVEAASGFTSEE-----KDPEEFNLIL--FHHIL-----RVEPL-LKIRS 697
Db 228 FKXWYSENEPTPFLPEQLRTYCHNDTEILLKALVFRHILINKITQGFVLPISCTTAS 287
QY 698 AGOKVQDCVFPYQ-----IFMEKNEKVGVTIQQLLSEWSFINSLKFAEAPSLIIQ 748
Db 288 ACNNIFKAQPMQEDQLAMVPELGVERNDRASVIAI-KYLDWRAKSEGDIDQHAGN----- 341
QY 749 MPFRG--KDPKLPKKIIPPSLEMLNITLLEDTPOCRICGGLAMYTECREDYDDPDISAKI 806
Db 342 ----GREKQWKFEK-----LDGWIESQQRCEIEVIGCYWHGCDRCF-KPD----- 380
QY 807 KOFC--KTCNTQVHLHPKRLNHNKYNPVSPLPKDLPDWDWRHCCIPCQNMELFAVLCIETSH 864
Db 381 EQLVDEKTCOELNEMTQDRLURLRE-----PDTQGR-----SLQVBEIWECEIIND 425
QY 865 YVAFVKYKDDSAWLFPDSMADRDGGQNGFNIPQVTPC-----PEVGEYLKMSLE 914

Db 426 QLK-----KNHEMKVFFDDLANERGLD-----PRLAYCGGRTGRLFAEASADEKISVF 476
Qy 915 DLHSL 919
Db 477 DIVSL 481
RESULT 6
PCT-US05-03344-1
; Sequence 1, Application PC/TUS0503344
; GENERAL INFORMATION:
; APPLICANT: Dana Farber Cancer Institute
; APPLICANT: D'Andrea, Alan D
; TITLE OF INVENTION: Method for Determination and Quantification of Radiation or Genot
; TITLE OF INVENTION: Exposure
; FILE REFERENCE: 7032/2082
; CURRENT APPLICATION NUMBER: PCT/US05/03344
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/540380
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1451
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-03344-1
Query Match 2.4%; Score 119.5; DB 1; Length 1451;
Best Local Similarity 18.3%; Pred. No. 4.7;
Matches 198; Conservative 141; Mismatches 376; Indels 367; Gaps 47;
Qy 3 SGLWSEKVTSPYWEERIFVLLLOE--CSVTDKQTKLLK-----VPKGS1-----46
Db 409 SGCIQELQSTF---SVHYLVKDMCSSILSLAQSLHSLDQSIISFGSLLYKYAFKFF 465
Qy 47 GOYIQRDSVGHRSIPSAKGNKQI--GLKILEQPHAVLFVDEVDVINEKFTLELL-----100
Db 466 DTYCQOEVEGVALVTHICSGNEAEVDALDVLE---LVVLNPSAMMNAVFQGIIDYLD 522
Qy 101 --AITNCEERFSL-----FKNRNLSKGLQIDVGPVKVQLRSAGEKFP-----GVVRF 148
Db 523 NISPOQIRKLFYVLSLAFSKQNEASSHIQDDMHLVIRKQLSSTVPKYKLGIGAVTMA 582
Qy 149 GPLLARTVSGIFFGVELLEBERGQGTGCVQKQLFO---C-----DEDC 192
Db 583 GIMAADRSSES-----PSLTQERANLSDEQCTQVTSLLQLVHSCRSQSPQASALYYDEFA 636
Qy 193 GFVALDKLE-----LIEDDDTALESYAGP-----GDTWQ 222
Db 637 NLIQHEKLPKALEWGHGHTICNDQFQAVVDVSCVPEGDFPPVPKALYGLBEYDTQDGA 696
Qy 223 VELPPLPINSRSLKGGETI--ESG-----SLGYF-----245
Db 697 INLLPFLSQDPKADGPGVTSQESGKLYSPCLAPYFRLRLCVERQHNGNLEEDIGLL 756
Qy 246 --TVIFCDVLPKGB-----262
Db 757 DCPIFUTDLEPGKLESMGAKESFMCSLIFLTNNFREIVNAFCQETSPENKGVLTREL 816
Qy 263 -----VGVDMDNPIGNWGDGFDGVLCSFACVESTILLHINDIIPESVT--305
Db 817 KHIVELQILLEKYLAVTPDPVPLGNF-----VETL-----DITPHVTAI 858
Qy 306 --QERRPPKAFMRSGVDGKGSSSHNPKATGTSDB--GNRRSELFTVLNGSSVDQSQOS 362
Db 859 SAKIRKGGIERKQKTGSKTSSDLSLSEKNESECPTSHRQQLNKEFTGKEKTSLLL 918
Qy 363 KSKNTWYID-----EVAEDPAKSLTISTDPRSSPP 394
Db 919 HNSHAFFRELDIEVFSILHCGLVTKFILDTEMTATEVVLQGPPELFLLEDLSOKLES 978

Qy 395 LQPPVNSLTTNREPHSLPFSLTWMPNTNGSIGHSPL--SLSAQSV---MEELNTAPVQES 450
Db 979 MLTPPI-----ARRVPEFLKNK---GSRNIFGSHLQORSQAIEVHCVEQLLT-----1021
Qy 451 PPLAMPNGN-----SHGLEVSGSLAEVKEN-----PPFYGVIRWTIG--Q 486
Db 1022 -PMCNHLENIHNYIOCLAAENHGVVDGPGVKQVEYHIMSSCYORLLQIFHGLFAMSGFSQ 1080
Qy 487 PPGLNEVLAGL-----ELEDECAGCTDGTFRGTRYFTCALKKALFVKL 529
Db 1081 PENQNLLYSALHVLSSRLKQGEHSQPLEBELLSSQSVHYLQNFHQSIQPSFOCAL---YLIRL 1137
Qy 530 KSCRPDSRFASLOPVSNQIERCNSLAFGGLYSEVVEENTPPKMEKEGLIMICKKGIQ 589
Db 1138 LMVILEKTASQAQ---NKEKTIASLA-RQFLCRVM-----PSGDKESNI-----SNDQL 1182
Qy 590 HYNSC-YLDSLFLCLFAPSSV---LDTVLLRPKKNQDVVEYSETOELLRTTEIVNPLRIY 644
Db 1183 HALLCIYLEHTESILKAEIEIAQVGPPELINSFKDASSSTFTLIRH-----TPVVF 1235
Qy 645 GYVCATKIMKRLKILEKVEAASGFTSEKDPREFL--NILFHHILRVEPLLKIRSAGKV 702
Db 1236 RVWMA---ELEKIVKIEPGTAADSQQIHEEKLLVWNAVRDPSILINLIKVFDSHPVL 1291
Qy 703 QDCY-FYOIFEMKNEKVGVPITQ-----OLLEWSFINSNLKFAEAPSCILLIOMP 750
Db 1292 HVCLKVGRLEFVEAFKQCMPLDLSFRKHREDVLSLLETFLDTRLLHLLCGHSHKHQDT 1351
Qy 751 RPKDFFKLFKIPPSILENITDLEDTPPQCRICGGLAMVECECY-----DDPDISAGK 805
Db 1352 RUTQHVPLKK--TLELLV-----CRVKAMLTNNCREAFWLGNLKNRDLQGE 1398
Qy 806 IK 807
Db 1399 IK 1400
RESULT 7
US-11-046-346-1
; Sequence 1, Application US/11046346
; GENERAL INFORMATION:
; APPLICANT: Dana Farber Cancer Institute
; APPLICANT: D'Andrea, Alan D
; TITLE OF INVENTION: Method for Determination and Quantification of Radiation or Genot
; TITLE OF INVENTION: Exposure
; FILE REFERENCE: 7032/2082
; CURRENT APPLICATION NUMBER: US/11/046,346
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/540380
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-046-346-1
Query Match 2.4%; Score 119.5; DB 7; Length 1451;
Best Local Similarity 18.3%; Pred. No. 4.7;
Matches 198; Conservative 141; Mismatches 376; Indels 367; Gaps 47;
Qy 3 SGLWSEKVTSPYWEERIFVLLLOE--CSVTDKQTKLLK-----VPKGS1-----46
Db 409 SGCIQELQSTF---SVHYLVKDMCSSILSLAQSLHSLDQSIISFGSLLYKYAFKFF 465
Qy 47 GOYIQRDSVGHRSIPSAKGNKQI--GLKILEQPHAVLFVDEVDVINEKFTLELL-----100
Db 466 DTYCQOEVEGVALVTHICSGNEAEVDALDVLE---LVVLNPSAMMNAVFQGIIDYLD 522
Qy 101 --AITNCEERFSL-----FKNRNLSKGLQIDVGPVKVQLRSAGEKFP-----GVVRF 148
Db 523 NISPOQIRKLFYVLSLAFSKQNEASSHIQDDMHLVIRKQLSSTVPKYKLGIGAVTMA 582
Qy 149 GPLLARTVSGIFFGVELLEBERGQGTGCVQKQLFO---C-----DEDC 192
Db 583 GIMAADRSSES-----PSLTQERANLSDEQCTQVTSLLQLVHSCRSQSPQASALYYDEFA 636
Qy 193 GFVALDKLE-----LIEDDDTALESYAGP-----GDTWQ 222
Db 637 NLIQHEKLPKALEWGHGHTICNDQFQAVVDVSCVPEGDFPPVPKALYGLBEYDTQDGA 696
Qy 223 VELPPLPINSRSLKGGETI--ESG-----SLGYF-----245
Db 697 INLLPFLSQDPKADGPGVTSQESGKLYSPCLAPYFRLRLCVERQHNGNLEEDIGLL 756
Qy 246 --TVIFCDVLPKGB-----262
Db 757 DCPIFUTDLEPGKLESMGAKESFMCSLIFLTNNFREIVNAFCQETSPENKGVLTREL 816
Qy 263 -----VGVDMDNPIGNWGDGFDGVLCSFACVESTILLHINDIIPESVT--305
Db 817 KHIVELQILLEKYLAVTPDPVPLGNF-----VETL-----DITPHVTAI 858
Qy 306 --QERRPPKAFMRSGVDGKGSSSHNPKATGTSDB--GNRRSELFTVLNGSSVDQSQOS 362
Db 859 SAKIRKGGIERKQKTGSKTSSDLSLSEKNESECPTSHRQQLNKEFTGKEKTSLLL 918
Qy 363 KSKNTWYID-----EVAEDPAKSLTISTDPRSSPP 394
Db 919 HNSHAFFRELDIEVFSILHCGLVTKFILDTEMTATEVVLQGPPELFLLEDLSOKLES 978

QY	149	GPLAERTVSGIIFFGVELLEEGRGQFTGDVYQKGQLFQ----	C-----	DEDC	192
Db		: : :			
Db	583	GIMAADRS-----PSLQERANLSDEQCTQVTSLLQLVHSCSEQSPQASALYDEFA	636		
QY	193	GFVALDKLE-----LIBDDDTALES DYAGP-----	GMTQ	222	
Db		: : :			
Db	637	NLIQHEKLDPKALEWVGHTICNDQDAFVVDSCVPEGDFFPVKALYLEEYDTQDGIA	696		
QY	223	VELPPELEINSRVSLKGETI--BSG-----		245	
Db		: : :			
Db	697	INLLPFLFSQDFAKDGGFVTSQSSGGKLVSPCLAPYFRLLRLCVRQHNGNLEEDGLL	756		
QY	246	--TWIFCDVLPJGKE-----SLGYP-----		262	
Db		: : :			
Db	757	DCPIFLDTLEPGKLEMSAKESAFMCSLFLFLTNWFEIWNAPCQTSPE NKGVTRL	816		
QY	263	-----VGVDMDNPIGNWGRFDGVLCSFACVESTILHINDIIPESVT--	305		
Db		: : :			
Db	817	KHIVELQILLEKYLA VTPDVVPLGNFD-----VEL-----	DITPHTVTAI	858	
QY	306	--QERRPPKLA FMSRGVDGKSSSHNKPATGSTDP--GNRRSELFTVLNGSSVDSPQS	362		
Db		: : :			
Db	859	SAKIRKGGKIERKQKTGSKTSSDTLSEBKNGECDTPSHRGQLNKEFTGKEKTSLLL	918		
QY	363	KSXNTWID-----EVAEDPAKSUTEISTEFDERSPP	394		
Db		: : :			
Db	919	HNSHAPFRELDIEVFSILHGLVTKFILDTEMHTEATEVQLGPPELLFLELDSQKLES	978		
QY	395	LQPPPVNSLTENRPHSLPSLTMPNWTNGSIGHSP--SLSAQSV--MEELNAPVOES	450		
Db		: : :			
Db	979	MLTPPI-----ARRVPELKNK-----GSRNIGFSLHQRSQAQEIHVCEBQLLT-----	1021		
QY	451	PPLAMPNGN-----SHGLEVGSLAEVKEN-----PPFYGVIRWIG--Q	486		
Db		: : :			
Db	1022	-PNCNHLNHNHYIOCLAEHNHGVDPGVKVQVEYHIMSSCYORLQIFHGLFANSGFSQ	1080		
QY	487	PPGLNEVLAGL-----ELEDEBCAGCTDGTFRGTRYFTCAKKALFVKL	529		
Db		: : :			
Db	1081	PENQNLLYSALHVLSSRLKOGESQPLELLSQSVHYLQNFHOSIPSFOCAL---YLRL	1137		
QY	530	KSCRPSRPFASLPQVSNQIERNCSLAPGGVLSVEVENTPPKMEKEGLEIMIGKKGIQG	589		
Db		: : :			
Db	1138	LMVILEKSTASAQ---NKEKIASLA-RQFLCRVW-----PSGDKEKSN-----SNDQL	1182		
QY	590	HYNSC-YLDSLTLECLFAPSV---LDTVLLRPKXNDVEYSETQELLATEIVNPLRIY	644		
Db		: : :			
Db	1183	HALLCTYLEHTESILKAIEBIAQVGPVELINSFKDASSSTFPTLTRH-----TPVVF	1235		
QY	645	GYVCATKIMKRLKILEVEAASGFTSEKDPBEFL--NILFHHILRVEPLLKIRSAGQV	702		
Db		: : :			
Db	1236	RVMA---ELEKIVKIEBPTAADSQOIHBEKLLYNMAVRDPSILINILKRVDFSHPV	1291		
QY	703	QDCY-EVQIFMEKNEKVGVP TIQ-----QLLEWSFINSNLKFAEPSCLIIQMP	750		
Db		: : :			
Db	1292	HVCLKVGRLFVEAFLEKQCMPLLDISFKKHREDVLSLETTQDLTRLHLHLCGHSKHQDT	1351		
QY	751	RFKQDFKLPFKIPPSLELNTDILLEDTPQRCICGGLAMYECCY-----DDPDISAGK	805		
Db		: : :			
Db	1352	RLTQHVP LKK---TLELV-----CRVKAMLTLANCREAFWLGNLKNRDQGE	1398		
QY	806	IK	807		
Db		: :			
Db	1399	IK	1400		

RESULT 8

RESUME 8
US-60-643-717-7896
; Sequence 7896, Application US/60643717
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for
; FILE REFERENCE: 38-21 (53629)A

```

, , CURRENT APPLICATION NUMBER: US/60/643,717
, ,
, , CURRENT FILING DATE: 2005-01-12
, ,
, , NUMBER OF SEQ ID NOS: 19247
, ,
, , SEQ ID NO 7896
, ,
, , LENGTH: 2412
, ,
, , TYPE: prt
, ,
, , ORGANISM: ASPERGILLUS NIDULANS FGSC A4
US-60-643,717-7896

```

Query Match	2.4%	Score 119.5	DB 8	Length 2412
Best local similarity	19.7%	Pred. No. 9.9		
Matches 113	Conservative 81	Mismatches 212	Indels 169	Gaps 28
Qy	330	KPKATGSTDGPNRRSE	-----LFYTIAGS-----SVDQSQSKNTWVI	-----370
Db	292	EPAREDTPSDGVALPEGALATLQFIPLEGTPFPWSETSEP	-----NBEEMQVAAPPERRRL	348
Qy	371	DEVAEDPAKSLTEIS	-----TDFDRSSPPLQPPVNSLTATENRFHS	411
Db	349	DD--ETFPVSELSSEYFTGLSIPSPGFFNSLAPRAHRTWSLPKLNQPP TSA--TAERFYN	404	
Qy	412	LPFSLTKMPTNTGSIHGSPLSLSAQSVMBELNTPVQESPP LA--MPPGNHSGLEVGSLA	469	
Db	405	LPFN-----REEGEITQEVIDLPERLNDEQLTAI---YAPPTAIKIPESPAHPPTGSI	456	
Qy	470	EYKENVPPFGVTRWICQP-----PGLNEVLAGELEDECACTDGT---	FRGTRYFTCAL	521
Db	457	PYSER-----VHEISRPTATYDPPDQDENYAEHLKHALSSLDRTSVMLAAQASYAAL	510	
Qy	522	KKALFVKL--KSCRPDSRFASLQPVNSNOIERCNSLAFGGYLSEVVEENTPPKWEKE---	575	
Db	511	RETNPVNNLPDSEDERQDVDESQHVSPALERNASVCFMGMPPEPPSSSLPAANASKDSIY	570	
Qy	576	--GLEIMIGKKK-----GIOGHYNSCYLSDTLFCL-----	F	604
Db	571	WRGFRELLDQSRSDRTFVHRSTFEDAVQSFRLGSLGNHKNCLLGNVELVLPDRPAYSGPF	630	
Qy	605	AFS-----SULDTVLLRKEKNVVEYSETQELRLTEI VNP---LRIYGYVCATKIM---	KL	655
Db	631	AKAPRHSVLPGLLQKAEFSMT---EKEQLVLSQISQPMWAMEALRYLQGGNLVVSPAR	686	
Qy	656	RKILEKVEAASGFTSEKDPESFNLFLH-----HILRVEPLLKIRSAQKQVQDCYFY	708	
Db	687	KFRSKRATAAAPHKTPKRQVRVLDLGGHATAEAWAHLAHDYPHVKV-----	Y	734
Qy	709	QIFMEKNKVGVPVTIQQLLEWFSFNSLNKFAEAPSCLIIQMRFPGDKFLFKKIFPSLEL	768	
Db	735	TYVTEHQQ-----VNKAIK--GPPNHRHIOVP-----QLWKLPFFDNKF	771	
Qy	769	NITD-----LLEDTPR--QCRICGGGLAMYECCREC	795	
Db	772	DVISARSLPAFLKTERPAGDCLDEYDCLCKECRCR	806	
RESULT 9				
US-10-450-763-51654				
; Sequence 51654, Application US/10450763				
; GENERAL INFORMATION:				
; APPLICANT: Hyseq, Inc				
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES				
; FILE REFERENCE: 790CIP3/US				
; CURRENT APPLICATION NUMBER: US/10/450,763				
; CURRENT FILING DATE: 2003-06-11				
; PRIOR APPLICATION NUMBER: PCT/US01/08631				
; PRIOR FILING DATE: 2001-03-30				
; PRIOR APPLICATION NUMBER: 09/540,217				
; PRIOR FILING DATE: 2000-03-31				
; PRIOR APPLICATION NUMBER: 09/649,167				
; PRIOR FILING DATE: 2000-08-23				
; NUMBER OF SEQ ID NOS: 60736				
; SOFTWARE: Custom				
; SEQ ID NO 51654				
LENGTH: 2543				

Db 268 SVSEDLLEFFIAVSAQAWNHQSGVVRPDSHQSGP-----NSDPGLEPED-----SNSTSA 318
QY 357 DSQP-----QSKSKNTWYIDEVADPAKSLTE-----ISTDFDRSSPPLQ- 396
Db 319 LEDPLEFLDMABTEIKICDYLFNVS DSSALNAKNGLTAKARDINAVLIDMERQGDVYRQ 378
QY 397 --PPVNSLTENR-----FHSLP-----FSLTKMPNTNGS-----IGH 428
Db 379 GTTPPIWHLTDKKRERMOIKRNTNSVPETAPAAIPETRRNAEFLTCNIPTSNASNNMVT 438
QY 429 SPLSLSAQSVMBELNAPVQESPPPLAMPNGSHGLEVGLAEVKENPPFVGVIRWIGQ-- 486
Db 439 EKVENQOEPIVKLENQEARPEARLKPVPVHYNGPSKAGVYDFENG-----QWATDDI 491
QY 487 PPGLEVLAGLELEDEACAGCTDGTFRGTRYFTCALKKALFKV--LKSCRDPDSRFASLQ-- 542
Db 492 PDDLNSIRA-----APGEFR-----AIME MPSFYSHGLPRCSYKYLTECOLK 534
QY 543 -PVSNOIERCNSLAFGYLSE--VVENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDST 599
Db 535 NPISGLLEYAQ---FASQTCFNMIEQSGPPHPRFKFQVINGREFPPPAEAGSKKVAQ 591
QY 600 LFCLEAFSSVLDVLLRPKEKN-DVEYVYSETQELLRT 635
Db 592 DAAMKANTILLBEAKAKDSKSEESHYSTEKESKT 628

RESULT 14

US-60-659-397-1214
; Sequence 1214, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERPERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1214
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1214

Query Match 2.2%; Score 113; DB 8; Length 1213;
Best Local Similarity 19.9%; Pred. No. 11;
Matches 115; Conservative 81; Mismatches 249; Indels 132; Gaps 26;
QY 142 PGVVRERGP-LLAERT-----VSGIFFGVELLEEGRGQGTGTVYQKOLFQCDSDCGF 194
Db 101 PG-LRPRFPVLLASSTRGRQVDIRGVRGVHLGSGQLGRGFQHPSPRGRSLPQRGVDCLS 159
QY 195 VALDKLELIEDDDTA-----LESYAGPGDTM-----QVELPPLLEINS---RVSLKGETI 242
Db 160 SHFOELSIYQDQQRILKLEELGEGKATTAHDLGKLTGPKKEINRVLSLAKKGLQK 219
QY 243 ESGTVIFCDVLPKESLGIVGVDMNDPIGNWDFDGVLCSPACVESTILLHNDIPIE 302
Db 220 EAGTPPLKIAVSTQAWNHQSGVVRP-----DGHSGQAPNSDPSLEP-----EDRNST 267
QY 303 SVTQERRPPKLAEMSR-----GVGDKGSSHNKPKATGSTDPGNRRSELFTYNGSSV 356
Db 268 SVSEDLLEFFIAVSAQAWNHQSGVVRPDSHQSGP-----NSDPGLEPED-----SNSTSA 318
QY 357 DSQP-----QSKSKNTWYIDEVADPAKSLTE-----ISTDFDRSSPPLQ- 396
Db 319 LEDPLEFLDMABTEIKICDYLFNVS DSSALNAKNGLTAKARDINAVLIDMERQGDVYRQ 378
QY 397 --PPVNSLTENR-----FHSLP-----FSLTKMPNTNGS-----IGH 428

Db 379 GTTPPIWHLTDKKRERMOIKRNTNSVPETAPAAIPETRRNAEFLTCNIPTSNASNNMVT 438
QY 429 SPLSLSAQSVMBELNAPVQESPPPLAMPNGSHGLEVGLAEVKENPPFVGVIRWIGQ-- 486
Db 439 EKVENQOEPIVKLENQEARPEARLKPVPVHYNGPSKAGVYDFENG-----QWATDDI 491
QY 487 PPGLEVLAGLELEDEACAGCTDGTFRGTRYFTCALKKALFKV--LKSCRDPDSRFASLQ-- 542
Db 492 PDDLNSIRA-----APGEFR-----AIME MPSFYSHGLPRCSYKYLTECOLK 534
QY 543 -PVSNOIERCNSLAFGYLSE--VVENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDST 599
Db 535 NPISGLLEYAQ---FASQTCFNMIEQSGPPHPRFKFQVINGREFPPPAEAGSKKVAQ 591
QY 600 LFCLEAFSSVLDVLLRPKEKN-DVEYVYSETQELLRT 635
Db 592 DAAMKANTILLBEAKAKDSKSEESHYSTEKESKT 628

RESULT 15

US-11-033-545-313
; Sequence 313, Application US/11033545
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/11/033,545
; CURRENT FILING DATE: 2005-01-12
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313
; LENGTH: 1226
; TYPE: PRT
; ORGANISM: Human
US-11-033-545-313

Query Match 2.2%; Score 113; DB 7; Length 1226;
Best Local Similarity 19.9%; Pred. No. 11;
Matches 115; Conservative 81; Mismatches 249; Indels 132; Gaps 26;
QY 142 PGVVRERGP-LLAERT-----VSGIFFGVELLEEGRGQGTGTVYQKOLFQCDSDCGF 194
Db 69 PG-LRPRFPVLLASSTRGRQVDIRGVRGVHLGSGQLGRGFQHPSPRGRSLPQRGVDCLS 127
QY 195 VALDKLELIEDDDTA-----LESYAGPGDTM-----QVELPPLLEINS---RVSLKGETI 242
Db 128 SHFOELSIYQDQQRILKLEELGEGKATTAHDLGKLTGPKKEINRVLSLAKKGLQK 187
QY 243 ESGTVIFCDVLPKESLGIVGVDMNDPIGNWDFDGVLCSPACVESTILLHNDIPIE 302
Db 188 EAGTPPLKIAVSTQAWNHQSGVVRP-----DGHSGQAPNSDPSLEP-----EDRNST 235
QY 303 SVTQERRPPKLAEMSR-----GVGDKGSSHNKPKATGSTDPGNRRSELFTYNGSSV 356
Db 236 SVSEDLLEFFIAVSAQAWNHQSGVVRPDSHQSGP-----NSDPGLEPED-----SNSTSA 286
QY 357 DSQP-----QSKSKNTWYIDEVADPAKSLTE-----ISTDFDRSSPPLQ- 396
Db 287 LEDPLEFLDMABTEIKICDYLFNVS DSSALNAKNGLTAKARDINAVLIDMERQGDVYRQ 346
QY 397 --PPVNSLTENR-----FHSLP-----FSLTKMPNTNGS-----IGH 428
Db 347 GTTPPIWHLTDKKRERMOIKRNTNSVPETAPAAIPETRRNAEFLTCNIPTSNASNNMVT 406
QY 429 SPLSLSAQSVMBELNAPVQESPPPLAMPNGSHGLEVGLAEVKENPPFVGVIRWIGQ-- 486
Db 407 EKVENQOEPIVKLENQEARPEARLKPVPVHYNGPSKAGVYDFENG-----QWATDDI 459

QY	487	PPG NEVL GLELE DE CAGCTD GTRGTR YFTC ALKK ALFVK--LK SCRPDS RFAS LQ--	542
Db	460	PDDLNSIRA-----APGEFR-----AIMEMPSYSHGLPRCSPYK L TECQLK	502
QY	543	-PVSNOIERCNSLAFGGYLSE--VVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDST	599
Db	503	NPISGLLEVAQ---FASQTCEFNMIQSGPPHPRFKFQV VINGRE FPFPAEAGSKK VAKQ	559
QY	600	LFCLFAPSSVLDTVLLRPKEKN-DVEYYSETQ ELLRT	635
Db	560	DAAMKAMTILLEEAKADSGKSESSHYS TEKESEKT	596

Search completed: April 18, 2005, 15:33:47
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:55:11 ; Search time 592 Seconds
(without alignments)
2496.870 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPWEERI.....RLLCDAYCMYQSPMTSLYK 949

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9157160 seqs, 778792062 residues

Total number of hits satisfying chosen parameters: 18314320

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09671687/runat 18042005 115020 17215/app query.fasta_1.1095
-DB=Pending Patents NA_New -QPMT=fastap -SUFFIX=rnnp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2o -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CEN 1 1 268 @runat 18042005 115020 17215 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents NA_New:

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2:
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq2:
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:
9: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2:
11: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3:
12: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4971.5	98.8	5371	7	US-10-760-678-3 Sequence 3, Appli
2	3676	73.0	2523	7	US-10-122-851-290 Sequence 290 App
C 3	1490.5	29.6	1160	8	US-10-450-763-26328 Sequence 26328, A
C 4	239	4.7	600	7	US-10-972-079-20538 Sequence 20538, A
C 5	234.5	4.7	600	7	US-10-972-079-20539 Sequence 20539, A
C 6	234.5	4.7	600	7	US-10-972-079-20540 Sequence 20540, A

7	160.5	3.2	4847	10	US-11-036-196-1764 Sequence 1764, Ap
8	151	3.0	5793	12	US-60-660-557-55 Sequence 55, Appl
9	151	3.0	5898	12	US-60-660-557-54 Sequence 54, Appl
10	146	2.9	5449	2	PCT-US05-07748-423 Sequence 423, App
11	146	2.9	5554	2	PCT-US05-07748-422 Sequence 422, App
12	146	2.9	5555	2	PCT-US05-07748-705 Sequence 705, App
13	136	2.7	2154	7	US-10-932-182A-77607 Sequence 77607, A
14	128	2.5	1497	8	US-10-450-763-21942 Sequence 21942, A
C 15	127	2.5	3918	7	US-10-467-657-5405 Sequence 5405, Ap
C 16	127	2.5	3918	7	US-10-467-657A-5405 Sequence 5405, Ap
17	126	2.5	4314	12	US-60-651-509-1122 Sequence 1122, Ap
18	126	2.5	4314	12	US-60-651-235-3001 Sequence 3001, Ap
19	126	2.5	4314	12	US-60-664-579-1592 Sequence 1592, Ap
20	126	2.5	6521	12	US-60-651-509-1123 Sequence 1123, Ap
21	126	2.5	6521	12	US-60-651-235-3002 Sequence 3002, Ap
22	126	2.5	6521	12	US-60-664-579-1593 Sequence 1593, Ap
C 23	125.5	2.5	4374	8	US-10-450-763-21946 Sequence 21946, A
24	123.5	2.5	4249	7	US-10-489-448-399 Sequence 399, App
25	123.5	2.5	5810	12	US-60-660-590-101 Sequence 101, App
26	123.5	2.5	5894	1	PCT-US05-10257-421 Sequence 421, App
27	123.5	2.5	7789	7	US-10-696-909A-45 Sequence 45, Appl
28	123.5	2.5	9027	7	US-10-696-909A-43 Sequence 43, Appl
29	123	2.4	1506	7	US-10-932-182A-2930 Sequence 2930, Ap
C 30	122.5	2.4	31826	10	US-11-031-175-1256 Sequence 1256, Ap
31	122	2.4	1500	7	US-10-932-182A-76457 Sequence 76457, A
C 32	120	2.4	2089378	7	US-10-526-324-1 Sequence 1, Appli
C 33	120	2.4	2089378	7	US-10-526-324-342 Sequence 342, App
C 34	120	2.4	2089378	7	US-10-526-324-723 Sequence 723, App
35	120	2.4	2089378	7	US-10-526-324-1087 Sequence 1087, Ap
36	120	2.4	2089378	7	US-10-526-324-1459 Sequence 1459, Ap
37	120	2.4	2089378	7	US-10-526-324-1838 Sequence 1838, Ap
C 38	119.5	2.4	5715	1	PCT-US05-10257-236 Sequence 236, App
C 39	119.5	2.4	9359	8	US-10-450-763-22954 Sequence 22954, A
40	119	2.4	6343	12	US-60-659-397-310 Sequence 310, App
41	119	2.4	6487	12	US-60-659-397-309 Sequence 309, App
42	119	2.4	11058	10	US-11-032-569-1 Sequence 1, Appli
43	118	2.3	2253	7	US-10-932-182A-79868 Sequence 79868, A
C 44	118	2.3	16047	10	US-11-031-175-1136 Sequence 1136, Ap
45	117.5	2.3	1828	2	PCT-US05-02325-51 Sequence 51, Appl

ALIGNMENTS

RESULT 1

US-10-760-678-3
; Sequence 3, Application US/10760678

; GENERAL INFORMATION:

; APPLICANT: Derry, Jonathan

; APPLICANT: Fanslow, William

; APPLICANT: Dougall, William

; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING

; FILE REFERENCE: 3198

; CURRENT APPLICATION NUMBER: US/10760,678

; CURRENT FILING DATE: 2004-01-20

; PRIOR APPLICATION NUMBER: US/09/851,673

; PRIOR FILING DATE: 2001-05-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 3

; LENGTH: 5371

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (392) .. (3262)

US-10-760-678-3

Alignment Scores:
Pred. No.: 0 Length: 5371
Score: 4971.50 Matches: 948
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 98.76% Indels: 7

DB:	7	Gaps:	5
US-09-671-687A-3 (1-949) x US-10-760-678-3 (1-5371)			
QY	1	MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle	20
DB	392	ATGAGTTTCAGGCTTATGGAGCCAGAAAAGATCACTTACCCCTACTGGGAAGAGCGGATT	451
QY	21	PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
DB	452	TTTTACTTCTCTTCAAGAATGCAGGTTACAGACAAACAAACAAACAAAGCTCCTTAA	511
QY	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle	60
DB	512	GTACCGAAGGGAAGTATAGACAGTATATTCAAGATCGTTCTGTGGGGCAATCAAGGATT	571
QY	61	ProSerAlaLysGlyLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80
DB	572	CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA	631
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99
DB	632	GTTCCTTTGTGATGAAAAGGATGTTGTAGAGATAATGAAAAGTTTCACAGATTACTT	691
QY	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
DB	692	TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCTGTTTAAAAACAGAAACAGACTTAAGT	751
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
DB	752	AAAGCGCTCCAAATAGACGTGGCTGCTCTGTGAAAGTACAGCTGAGATTCTGGGAAGAA	811
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
DB	812	AAATTTCTGGAGTTGTAGCTTCAGAGACCCCTGTTAGCAGAGAGACAGTCTCCGGA	871
QY	160	IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal	179
DB	872	ATATTTCTTTGGAGTTGAAATTGCTGGAAGAGGTCTGTGGTCAAGGTTTCACTGACGGGGT	931
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
DB	932	TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC	991
QY	199	LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218
DB	992	AAGCTAGAACTCATAGAAGATGATGACCTGCAATGGAAGATGATTACCAGAGTCTCTGG	1051
QY	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238
DB	1052	GACACAATGCAGGTCGAACTTCTCTCTTTGGAAATAAACTCCAGAGTTCCTTTGAAGTT	1111
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258
DB	1112	CGAGAAACAATAAGAACTCGAAACAGTTATATTCTGTGATGTTTGGCAGAAAGAAAGC	1171
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278
DB	1172	TTAGGATATTTTGTGGTGTGGACATGGAATAACCTATTGGCAACTGGGATGGAAGATT	1231
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297
DB	1232	GATGGAGTCAGCTTTGTAGTTTGTGTGTGTTGAAAGTACAATTTCTATTGGACATCAAT	1291
QY	298	AspIleIlePro-----GluSerValThrGlnGluArgProProLysLeuAla	314
DB	1292	GATATCATCCAGCTTTATCAGAGAGGTGTGACGACGAAAGGAGGCTCCCAAACCTGCC	1351
QY	315	PheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThr	334
DB	1352	TTTATGTCAAGAGGTGTGGGCAAGGTTTCATCCAGTCATATAATAACAAAGGCTACA	1411
QY	335	GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly	353
DB	1412	GGATCTACCTCAGACCCCTGGAATAGAAAACAGATCTGAATATTATTTATACCTTAAATGGG	1471
QY	354	SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal	373
DB	1472	TCTTCTGTGTGACTCACACACACATCCAAATCAAAAATACATGTTGATGAAGTT	1531
QY	374	AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerPro	393
DB	1532	GCAGAAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTCACCGCTTCTACCA	1591
QY	394	ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro	413
DB	1592	CCACTCCAGCCCTCCTCTGTGAACCTCACTGACACACAGAACAGATTCCACTCTTTACCA	1651
QY	414	PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu	433
DB	1652	TTCACTCTCACCAAGATGCCAATACCAATGGAAATATTGGCCACAGTCCACTTCTCTG	1711
QY	434	SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeu	453
DB	1712	TCAGCCCACTGTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCCCTG	1771
QY	454	AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu	473
DB	1772	GCCATGCCCTCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAG	1831
QY	474	AsnProProPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluVal	493
DB	1832	AACCTCTCTTCTATGGGGTAATCCGTTGGATCGGTGAGCCACAGAGACTGAATGAAGTG	1891
QY	494	LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly	513
DB	1892	CTCGCTGGACTGGAACCTGGAAGATGAGTGTGACGGCTGTACGGATGGAACTTCAGAGGC	1951
QY	514	ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg	533
DB	1952	ACTCGGTAATTTACCTGTGCTTGAAGAAGCGCTGTTTGTGAAACTGAAGAGCTGCAGG	2011
QY	534	ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer	553
DB	2012	CCTGACTCTAGGTTTCGATTCATTCAGCCGGTTTCCAATCAGATTGAGCGCTGTAACCT	2071
QY	554	LeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu	573
DB	2072	TTAGCAATTTGGAGCTTACTTTAGTGAAGTAGAAGAAATACTCCACCAAAATCGAA	2131
QY	574	LysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSer	593
DB	2132	AAAGAAGGCTTGGAGATAATGATTGGGAAGAAAGGCATCCAGGGTCAATCAATCT	2191
QY	594	CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal	613
DB	2192	TGTTACTTAGACTCAACCTTATTTCTGCTTATTTTGTCTTTTAGTTCTGTGTCACACTGTG	2251
QY	614	LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeu	633
DB	2252	TTACTTAGACCCCAAGAAAGAACGATGAGAATATTATAGTGAAACCCAGAGACTACTG	2311
QY	634	ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet	653
DB	2312	AGGACAGAAATTCGTTAATCCTCTGAGATAATATGGATATGTGTGTCACAAAAATATG	2371
QY	654	LysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLys	673
DB	2372	AAACTGAGGAAAAATCTTGAAGAGTGGAGGCTGCATCAGGATTTTACCTCTCGAAGAAAA	2431
QY	674	AspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeu	693
DB	2432	GATCCTGAGGAATCTTGAAATATTTCTGTTTCATCATATTTTAAGGGTAGAACCTTTGCTA	2491
QY	694	LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu	713

```
Db 2492 AAAATAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTTACTTCTATCAAAATTTTATGGAA 2551
Qy 714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn 733
Db 2552 AAAAATCAGAAATGCGGTGCCCAATTCAGCAGTTGTAGAAATGGTCTTTTATCAAC 2611
Qy 734 SerAsnLeuLysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGly 753
Db 2612 AGTAACCTGAAATTTGAGAGGACCACATCATGTGATTATTCAGATGCGCTCGATTGGA 2671
Qy 754 LysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
Db 2672 AAAGACTTTAAACTATTAAAAAATTTTCTCTCTGGAATTAATAACAGATTTA 2731
Qy 774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
Db 2732 CTTGAAGACACTCCAGACAGTCCGATATGTGGAGGGCTTGCAATGTATGAGTGAGA 2791
Qy 794 GluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCys 813
Db 2792 GAATGCTAGCAGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTGTAACCTGC 2851
Qy 814 AsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeu 833
Db 2852 AACACTCAAGTCCACCTTCATCCGAGAGGCTGATCATTAATATTAACCCAGTGTCACTT 2911
Qy 834 ProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGlu 853
Db 2912 CCCAAAGACTTACCAGCTGGGACTGGAGACACGGCTGCATCCCTTGGCCAGATATGGAG 2971
Qy 854 LeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLys 873
Db 2972 TTATTTGCTGTTCTCTGCATAGAAACAAAGCCACTATGTTGCTTTTGTGAAAGTAGTGG 3031
Qy 874 AspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGly 893
Db 3032 GACGATTCGCCCTGGCTCTCTTTGACAGCATGGCCGATCGGGATGGTGTGAGATGGC 3091
Qy 894 PheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeu 913
Db 3092 TTCAACATTCTCAAGTCACTCCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCCTG 3151
Qy 914 GluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCys 933
Db 3152 GAAGACCTGCATTCCTTGGACTCCAGGAGAAATCCAGGCTGTGACGAGACTGCTTTGT 3211
Qy 934 AspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3212 GATGCATATATGTGTCATGTACAGAGTCCAAACAATGAGTTTGTACAAA 3259
```

RESULT 2

US-10-122-851-290
; Sequence 290, Application US/10122851

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyang
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDV3
; CURRENT FILING DATE: US/10/122,851
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 290

; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2138)
US-10-122-851-290
Alignment Scores:
Pred. No.: 3,03e-242 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 7 Gaps: 2
US-09-671-687A-3 (1-949) x US-10-122-851-290 (1-2523)

```
Qy 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
Db 45 GGAATAATAAAGCTTTTATATATTTTGGTGTGGACATGGATAACCCCTATTGGCAACTGG 104
Qy 275 AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu 293
Db 105 GATGGAAGATTGATGGAGTGCAGCTTTGTAGTTTTGCGTGTGTGAAAGTACAATTCTA 164
Qy 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeu 313
Db 165 TTGCACATCAATGATATATCATCCAGAGAGTGTGACGAGAAAGAGAGGCTCCCAAACTT 224
Qy 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333
Db 225 GCCTTTATGTCAAGAGGTGTTGGGACAAAGGTTTCATCCAGTCAATAATAAACCAAGGCT 284
Qy 334 ThrGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352
Db 285 ACAGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAATTTATTTTATACCTTAAAT 344
Qy 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGlu 372
Db 345 GGGTCTCTGTTGACTCACAACCAACATCCAAATCAAAAATAACATGATGATGATA 404
Qy 373 ValAlaGluAspProAlaLysSerLeuThrGluLysSerThrAspPheAspArgSerSer 392
Db 405 GTTGCAAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTTGCCTTCTTCA 464
Qy 393 ProProLeuGlnProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeu 412
Db 465 CCACCACTCCAGGCTCTCTCTGTGAACTCACTGACCCAGAGAAACAGATTCCACTCTTTA 524
Qy 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
Db 525 CCATTTCAGTCTCACAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCT 584
Qy 433 LeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProPro 452
Db 585 CTGTCAGCCAGCTGTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCC 644
Qy 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
Db 645 TTGGCCATGCTCTCTGGGAACTCACAATGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAG 704
Qy 473 GluAsnProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
Db 705 GAGAACCTCTCTTCTATGGGGTAAATCCGTGGAGTCCGTGAGCCACAGGACTGTAATGAA 764
Qy 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
Db 765 GTGCTCGCTGGACTGGAACTGGAAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAG 824
Qy 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys 532
Db 825 GGCACCTGGTATTTTACCTTGCCCTCGAAGAGCGCTGTTTGTGAAACTGAAGAGCTGC 884
```

```
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552
DB 885 AGGCGTGACTCTAGAGTTGCATCATTTGCAGCGGTTTCCAAATCAGATTGAGCGCTGTAAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMet 572
DB 945 TCTTTAGCATTTGGAGGCTTACTTAAGTGAAGTAGTAGAAGAAATACTCCACAAAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAGAAGGCTTGGAGATATGATTGGAGAGAGAAAGGCATCCAGGGTCATTACAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTTACTTAGACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTTACTTAGACCCCAAGAAAGACGATGTAGAATATTATAGTGAAACCCCAAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACGAAATTTGTTAATCTCTGAGAAATATATGGATATGTGTGCCACAAAAT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATGAAACTGAGGAAATACTTCAAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAA 1304
QY 673 LysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCTCTGAGGAATTTCTTGAATATCTTGTTCATCATATTTTAAGGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAAAAATAGATACGACAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCGTTCCCAATTCAGCAGTTGTAGATGGTCTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTGCAGAGGCACCATCATGTCTGATTATTCAGATGCCCTGATTT 1544
QY 753 GlyLysAspPheLysLysLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GAAAAAGACTTTAAACATATTTAAAAAATTTTCTCTCTGGAATTTAAATATAACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTTGAAGACACTCCAGACAGTCCCGATATGTGGAGGGCTTGCAATGTATAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAGATGCTACGACCATCGGACATCTCAGCTGGAAAAAATCAAGCAGTTTGTGAAAAACC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCAATCCGAAGAGGCTGAATCATAAATATAACCCAGGTGCA 1784
QY 833 LeuProLysAspLeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGAGACTTACCCGACTGGGACTGGAGACAGCGCTGCATCTCTTCCAGAAATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTTATTTGCTGTTCTGCTAGAAACAAAGCCACTATGTTGCTTTTGTGAAGTATGGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGGACGATTTCTCGCTGGCTCTTTTGACAGCATGCCCGATCGGGATGTGTGTGCAAT 1964
```

```
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912
DB 1965 GGCTTCAACATTCCTCAAGTCACCCCATGCCCAGAAAGTAGGAGAGTACTTTGAAGATGCT 2024
QY 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
DB 2025 CTGGAAGACTGTCATTCCTTGGACTCCAGGAGAAATCCAGGCTGTGACCGAAGACTGCTT 2084
QY 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 2085 TGTGATCATATATGTGTCATGTACCAAGAGTCCCAACAATGAGTTGTGTACAAA 2135

RESULT 3
US-10-450-763-26328/c
; Sequence 26328, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hveeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 26328
; LENGTH: 1160
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (639)..(67)
; OTHER INFORMATION: 96% homologous to Homo sapiens Familial Cylindromatosis
; OTHER INFORMATION: Gene, accession number AJ250014, Smith-Waterman Score=963.
US-10-450-763-26328

Alignment Scores:
Pred. No.: 1,09e-92 Length: 1160
Score: 1490.50 Matches: 295
Percent Similarity: 97.38% Conservative: 2
Best Local Similarity: 96.72% Mismatches: 4
Query Match: 29.61% Indels: 4
DB: 8 Gaps: 3

US-09-671-687A-3 (1-949) x US-10-450-763-26328 (1-1160)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 979 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAGTACATTCACCTACTCGGAGAGCGGATT 920
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 919 TTTTACTTGTCTTCAAGAAATGCAGGTTACAGACAAACAAACACAAAGCTCTCTTAAA 860
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHiserArgIle 60
DB 859 GTACCCGAAGGAAAGTAGTAGACAGATATATTTCAAGATCGTTCTGTGGGCGATTCAAGGATT 800
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 799 CCTTCTGCAAAAGGCAAGAAATAATCAGATTGGATTAAAAATTTAGAGAACACCTCATGCA 740
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
DB 739 GTTCTCTTGTGATGAAAAGGATGTTGTAGAGATAAATGATATAAGTTCACAGAGTTACTT 680
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArg-LeuSe 119
DB 679 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCCTGTTTAAACCCAGAAACCCAGTTTAAG 620
```

QY 119 rLysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGI 139
Db 619 TAAAGCCCTCCAAATAGACGGGGCTGTCTCTGTGAAAGTACACTGAGATCTCGGGAAGA 560
QY 139 uLysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGI 159
Db 559 AAAATTCTCGAGTGTGACCTCTCAGAGGACCCCTGTAGCAGAGGACGACGTCTCCGG 500
QY 159 YLePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlnPheThrAspGlyVa 179
Db 499 AATATTCTTTGGAGTTGAATTGCTGGAAGAAGGTCGTGGTCAAGGTTTCACTGACGGGT 440
QY 179 lTyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAs 198
Db 439 GTACCAAGGGAACACACTTTTTCAGTGTGTGATGAAGATTGCGCGTGTGTGTCATTGGA 380
QY 198 pLysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGI 218
Db 379 CAAGCTAGAACTCATAGAAGATGACACTGCAATTGGAATGATTTACGAGGTCCTGG 320
QY 218 YAspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGI 238
Db 319 GGACACAATGCAGGTGCACTTCTCTCTTTGGAATAAACTCCAGAGTTCTTTGAAGGT 260
QY 238 YGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258
Db 259 TGGAGAAACAATAGAACTCTGGAACAGTTATATCTGTGATGTTTTCGACGAAAGAAAG 200
QY 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyValAsnTrpAspGlyArgPh 278
Db 199 CTTAGCATATTTGTTGGTGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAAGATT 140
QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297
Db 139 TGATGGAGTGCAGCTTTGTAGTTTTCGCTGTGTGTTGAAGTACAAATCTATTGACATCAA 80
QY 297 nAspIleIlePro 301
Db 79 TGATATCATCCCCA 67
RESULT 4
US-10-972-079-20538/c
; Sequence 20538, Application US/10972079
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; PRIOR FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20538
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894231145_1
US-10-972-079-20538
Alignment Scores:
Pred. No.: 3,69e-07 Length: 600
Score: 239.00 Matches: 66
Percent Similarity: 50.64% Conservativeness: 13
Best Local Similarity: 42.31% Mismatches: 47
Query Match: 4.75% Indels: 30
DB: 7 Gaps: 7

US-09-671-687A-3 (1-949) x US-10-972-079-20538 (1-600)
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsn 715
Db 450 AGATCAGCAGGTCAAGAAATGCAAGACTGTTATTTCTACCAATTTTATGACACAAAAT 391
QY 716 GluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsn 735
Db 390 GAGAAAGTTGGAGTCCCGACAATYCACAGTTACTGGAGTGTCTCTTCATCAACAGCAAC 331
QY 736 LeuLysPheAlaGlu---AlaProSerCysLeuIleIleGlnMetProArgPheGlyLys 754
Db 330 TTGAAGTTTCAGAGGTTGGTGACACTGKGTGTGTCT---AGAAATACTGAT 280
QY 755 AspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeu 774
Db 279 AGCTACATCTGTTTCAGC-----AATGAGAGAAAGTGGGA 244
QY 775 GluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGlu 794
Db 243 TTGGGTGTGAAAGGAAATGCAGA---TGTGTTGGGATGGTGGAGTGTCTCAGTG----- 193
QY 795 CysTyrAspAspProaspIleSerAlaGlyLysIleLysGlnPheCys----- 810
Db 192 -----LysTyrCysAsnThrGlnValHisLeuHisProLysArgLeu-----Asn 151
QY 811 -----LysTyrCysAsnThrGlnValHisLeuHisProLysArgLeu-----Asn 825
Db 150 GGTGTGGGTGTACCAAAATGCTTTCATCTATTAGCCAGGAATTTACTCAATAAATAC 91
QY 826 HisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAsp 841
Db 90 TCTTCGTTACTGTGTGAAGTTGCGCTTAAGATCTAGTATTTCGGAT 43
RESULT 5
US-10-972-079-20539/c
; Sequence 20539, Application US/10972079
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; PRIOR FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20539
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894231145_2
US-10-972-079-20539
Alignment Scores:
Pred. No.: 7,51e-07 Length: 600
Score: 234.50 Matches: 45
Percent Similarity: 90.74% Conservativeness: 4
Best Local Similarity: 83.33% Mismatches: 4
Query Match: 4.66% Indels: 1
DB: 7 Gaps: 1
US-09-671-687A-3 (1-949) x US-10-972-079-20539 (1-600)
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsn 715
Db 383 AGATCAGCAGGTCAAGAAATGCAAGACTGTTATTTCTACCAATTTTATGACACAAAAT 324


```
QY 426 IleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAasnThrAla 445
Db 758 GACGGCCACTCAGTGAATCCCTCACCAGCCAGAACTGTCCTGCATTCCTGGCAGCGCC 817
QY 446 ProValGlnGluSerProProLeu----- 453
Db 818 -----ACGCCCCCACTCAGCGCCGAGTATCCCTCGGGAGAGTGTCTC 865
QY 454 -----AlaMetProProGlyAasnSerHisGlyLeuGluValGlySerLeuAlaGluVal 471
Db 866 AACAGCTCCGTGAAGACGGGCAATGATCTGGCTCCAACTCTCGGACAGTGTCTGTG 925
QY 472 LysGluAasnProProPheTyr----- 478
Db 926 AAGCGTGTGACAGGACCTCCACCTGGGAGACCGTGTGCTGTGGTGGACCAAGACC 985
QY 479 GlyValIleArgTrpIleGlyGlnProProGlyLeuAasnGluValLeuAlaGlyLeuGlu 498
Db 986 GGTGTGTAGATATGTTGGGAGACGACTTGGCCAAAGCGAGTGGTGTGTGTGGAG 1045
QY 499 LeuGluAaspGluCysAlaGlyCysThrAaspGlyThrPheArgGlyThrArgTyrPheThr 518
Db 1046 CTG---GACGAGCCCTTGGGAAGAAGCATGGGCGAGTGGCAGGACCCAGGTACTTCCAG 1102
QY 519 CysAlaLeuLysLysAlaLeuPheValLysLeu----- 529
Db 1103 TGCCCAACCAGTTCGTCTCTTTGCAACCAATCCACAAGGTCATCCGAATTTGGCTTCCCA 1162
QY 530 -----LysSerCysArgProAaspSerArgPheAla----- 539
Db 1163 TCTACAGTCCAGCCCAAGGCCAAGAGACCAAGCGATGCTATGGGTGTCTCAGCGCTTG 1222
QY 540 SerLeuGlnProValSerAasnGlnIleGluArgCysAasnSerLeuAla-----PheGly 557
Db 1223 ACCCAGCCCGCAGCAGTCTTCCATCAGCTCTGTGAGTCTGTGCGCTCTCTGTGTGT 1282
QY 558 GlyTyrLeuSerGlu-----ValValGluGluAasnThrProProLysMetGluLys 574
Db 1283 GGCAGCGCCGAGCGGTAGTGCCTGTCTCACAGAGACCTCTTACGCTATGCCCGGAAGATC 1342
QY 575 GluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyr-AasnSerCy 594
Db 1343 TCAGGCACACAGCCCTCGAGGAGGCATCGAAGGAGAAACAGCAGCAGCAGCATCGAAGCCTG 1402
QY 594 sTyrLeuAaspSerThr 599
Db 1403 CTGGCTGACGCTGACT 1418
```

RESULT 8

```
US-60-660-557-55
; Sequence 55, Application US/60660557
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Ming-Yi Chiang
; APPLICANT: Nicholas M. Dean
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Susan M. Freier
; APPLICANT: Ravi Jain
; APPLICANT: Donna T. Ward
; TITLE OF INVENTION: COMPOSITIONS AND THEIR USES DIRECTED TO ORGANELLE COMPONENTS
; FILE REFERENCE: DPTK-0075US.L
; CURRENT APPLICATION NUMBER: US/60/660,557
; CURRENT FILING DATE: 2005-03-10
; NUMBER OF SEQ ID NOS: 1853
; SOFTWARE: PatentSeq version 1.0
; SEQ ID NO 55
; LENGTH: 5793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-660-557-55
Alignment Scores:
```

```
Pred. No.: 6 69 Length: 5793
Score: 151.00 Matches: 107
Percent Similarity: 31.58% Conservative: 61
Best Local Similarity: 20.11% Mismatches: 184
Query Match: 3.00% Indels: 180
DB: 12 Gaps: 23
```

US-09-671-687A-3 (1-949) x US-60-660-557-55 (1-5793)

```
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAasnSerArgValSerLeuLysGly 238
Db 300 GAGACTCAGGAGGAATTTGTGGATGACTTTTCGAGTTGGGAGCGAGTTTGGTGATGCA 359
QY 239 -----GlyGluThrIleGluSerGlyThrValIlePhe 249
Db 360 AATAAGCCTGGATTTATCCAGTTTCTTGGAGAAACC----- 395
QY 250 CysAaspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAaspMetAaspAen 269
Db 396 ---CAGTTTCACAGGCCAG-----TGGGCTGGAATTTGTTTAGATGAA 437
QY 270 ProIleGlyAasnTrpAaspGlyArgPheAaspGlyValLeuCysSerPheAlaCysValGlu 289
Db 438 CCATAGCGAAGAACGATGTTCCGTGGCAGGAGTTCGG---TATTTCCAGTGTCAACCT 494
QY 290 SerThrIleLeuLeuHisIleAasnAaspIleIleProGluSerValThrGlnGluArgArg 309
Db 495 -----TTAAAGGGCATATTT-----ACCCGA 515
QY 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAaspLysGlySerSerHisAasn 329
Db 516 CCTTCAAAAGTTA-----ACAAGGAAGTG-----CAAGCA 545
QY 330 LysProLysAlaThrGlySerThrSerAaspProGlyAasnArgArgSerGluLeuPheTyr 349
Db 546 GAGATGAACCTAATGGCTGCAGACAAAGCCGCCCTCCCGAGCTACTTCCCGCTGTC 605
QY 350 ThrLeuAasnGlySerSerValAaspSerGlnProGlnSerLysSerLysAasnThrTrpTyr 369
Db 606 ACTTCTACGGCCAGCAGTGTGTCTCTCCCTCCACCCCTTCAAAAC----- 653
QY 370 IleAaspGluValAlaGluAaspProAlaLysSerLeuThrGluIleSerThrAaspPheAsp 389
Db 654 ATCCCTCAGAAACCATCACAGCCAGCA----- 680
QY 390 ArgSerSerProProLeuGlnProProProValAasnSerLeuThrThrGluAasnArgPhe 409
Db 681 ---GCAAGGAACCTTCAGCTACGCTCCGATCAGCAACCTTACA----- 722
QY 410 HisSerLeuProPheSerLeuThrLysMetProAasnThrAasnGlySerIleGlyHisSer 429
Db 722 ----- 722
QY 430 ProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAasnThrAlaProValGlnGlu 449
Db 723 -----AAAACTGCCAGTGAATCTATCTCAACCTTTTCAGAGGTGGCTCAATCAAGAAA 776
QY 450 SerProLeuAlaMetProProGlyAasnSerHisGlyLeuGluValGlySerLeuAla 469
Db 777 -----GGAGAAAGAGAGCTCAAAATCGGAGACAGATA 809
QY 470 GluValLysGluAasnProProPheTyrGlyValIleArgTrpIleGlyGlnProGly 489
Db 810 TTGTTGGTGGCAGTAAAGCT-----GGTGTAGTCGGTTTTCTTGGGAGACCGACTTT 863
QY 490 LeuAasnGluValLeuAlaGlyLeuGluLeuAaspGluCysAlaGlyCysThrAaspGly 509
Db 864 GCCAAGGGGAGTGGTGTGCGTGGAGTTA---GATGAGCCACTTGGGAAGAAATGATGCG 920
QY 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
Db 921 GCTGTTGCTGGAACAAGGTATTTTCAGTGTCAACCCAAATATGCGCTTGTTCGCT----- 974
```

QY 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
DB 975 -----CCTGTCAC 983
QY 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGlnAsnThrPro 569
DB 984 AAGTTACCAAGATTGGCTTC-----CCTCCACTACCA 1019
QY 570 ProLysMetGluLysGluGlyLeu---GluIleMetIleGlyLysLysLysGlyIleGln 588
DB 1020 GCCAAGCCCAAGCCCAAGCAGTGAGCGGAGTGATGGCCAGCACGTCGCCAGCCTGAAG 1079
QY 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
DB 1080 CGCAGCCCTTCTGCTCTTCCCTCAGCTCC-----ATGAGTCA 1118
QY 609 ValIleAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
DB 1119 GTGGCTCTCTGTGAGCAGCAGCC-----AGTCGG 1151
QY 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
DB 1152 ACAGGACTATTGACTGAACCTCTCCGTTAGCCAGGAGATCTCCGTTACCACTGCC 1211
QY 649 AlaThrLysIleMetLys-----LeuArgLysIle 658
DB 1212 CTCAGGAGGCCCTGAAGGAGAGCAGCAGCATTGAGCAGCTGTGGCGGAACGGAT 1271
QY 659 LeuGluValValGluAlaAlaSerGlyPheThrSerGluLysAspProGluGluPhe 678
DB 1272 CTGGAGAGCGGAGGTGGCCAGGCC---ACAGGCCACGTGGGGAGATAGAGCAGGAG 1328
QY 679 LeuAsnIleLeu-----PheHisIleLeu 687
DB 1329 CTAGCTCTGGCCCGGACGAGATGACCAGCATGTCTGGAATGGAGCCAAATGGAC 1388
QY 688 -----ArgValGluProLeuLysIle 695
DB 1389 CAGCTCGCAACAATGTGGAGCTGTGACAGGAGAGGTGGAGCTTCTCAACCAGCTT 1448
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPhe 707
DB 1449 GAAGAGGAGAAAGGAAGGTGGAGCCTTCAGTTC 1484

RESULT 9

US-60-660-557-54
; Sequence 54, Application US/60660557
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Ming-Yi Chiang
; APPLICANT: Nicholas M. Dean
; APPLICANT: Kenneth W. Doble
; APPLICANT: Susan M. Freier
; APPLICANT: Ravi Jain
; APPLICANT: Donna T. Ward
; TITLE OF INVENTION: COMPOSITIONS AND THEIR USES DIRECTED TO ORGANELLE COMPONENTS
; FILE REFERENCE: DPTK-0075US.L
; CURRENT APPLICATION NUMBER: US/60/660,557
; CURRENT FILING DATE: 2005-03-10
; NUMBER OF SEQ ID NOS: 1853
; SOFTWARE: PatentSeq version 1.0
; SEQ ID NO 54
; LENGTH: 5898
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-660-557-54

Alignment Scores:

Pred. No.:	6.84	Length:	5898
Score:	151.00	Matches:	107
Percent Similarity:	31.58%	Conservative:	61
Best Local Similarity:	20.11%	Mismatches:	184
Query Match:	3.00%	Indels:	180

DB: 12 Gaps: 23
US-09-671-687A-3 (1-949) x US-60-660-557-54 (1-5898)
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 300 GAGACTCAGGAGGAATTTGGGATGACTTTCGAGTTGGGAGCGAGTTTGGTGAATGGA 359
QY 239 -----GlyGluThrIleGluSerGlyThrValIlePhe 249
DB 360 AATAAGCCTGGATTATCCAGTTTCTTGGAGAAACC----- 395
QY 250 CysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsn 269
DB 396 ---CAGTTTGCACAGGCCAG-----TGGGCTGGAATTTGTTTATAGTAA 437
QY 270 ProIleGlyAsnTrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu 289
DB 438 CCATAGGCAAGAACCATGTTTCGGTGGCAGGAGTTCGG---TATTTCCAGTGTGAACCT 494
QY 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArg 309
DB 495 -----TTAAGGGCATATTT-----ACCGA 515
QY 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsn 329
DB 516 CTTCAAAAGTTA-----ACAAGGAGGTG-----CAAGCA 545
QY 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyr 349
DB 546 GAAGATGAAGCTAATGGCTGCAGACAAACCCGCTCCCGAGCTACTTCAACCGTGTGC 605
QY 350 ThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTyr 369
DB 606 ACTTCTACGCCAGCATGGTGTCTTCTCCCTCCACCCCTTCAAC----- 653
QY 370 IleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAsp 389
DB 654 ATCCCTCAGAAACCATCACAGCCAGCA----- 680
QY 390 ArgSerSerProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPhe 409
DB 681 ---GCAAAAGGAACCTTCAGCTACGCTCCGATCAGCAACCTTACA----- 722
QY 410 HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleHisSer 429
DB 722 ----- 722
QY 430 ProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGlu 449
DB 723 -----AAAACGCGCAGTGAATCTATCTCCAACCTTTTCAGAGGTGGCTCAATCAAGAAA 776
QY 450 SerProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAla 469
DB 777 -----GGAGAAAGAGAGCTCAAAATCGGAGACAGAGTA 809
QY 470 GluValLysGluAsnProProPheTyrValIleArgTrpIleGlyGlnProGly 489
DB 810 TTGGTTGGTGGCATAAGGCT-----GGTGTAGTCGGTTTCTTGGGAGACCGCACTTT 863
QY 490 LeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGly 509
DB 864 GCCAAGGGGAGTGGTGGGTGGAGTTA---GATGAGCCACTGGGAAGAATGATGCG 920
QY 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
DB 921 GCTGTTGCTGGAACAGAGTATTTTCAGTGTCAACCCAAATATGGCTTGTTCGCT----- 974
QY 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
DB 975 -----CCTGTCCAC 983
QY 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrPro 569

```
Db 984 AAAGTTTACCAAGATTGCTTC-----CCTTCCACTACACCA 1019
Qy 570 ProLysMetGluLysGluGlyLeu---GluLeuMetIleGlyLysLysGlyIleGln 588
Db 1020 GCCAAAGCCCAAGCCCAACGAGTGAGCGGAGTATGCGCCAGCAGCTGAAG 1079
Qy 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
Db 1080 CGCAGCCCTTCGCTCCCTTCCTCCCTCAGCTCC-----ATGAGCTCA 1118
Qy 609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
Db 1119 GTGGCTCCTCTGTGAGCAGCAGGCC-----AGTCGG 1151
Qy 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
Db 1152 ACAGGACTATTGACTGAACCTCTCCCGTTAGCCAGGAAGATCTCCGCTACCACTGCC 1211
Qy 649 AlaThrLysIleMetLys-----LeuArgLysIle 658
Db 1212 CTCAGGAGGCCCTGAAGGAGAGACAGCAGCAGCATTGAGCAGTGTGCGGGAACGGAT 1271
Qy 659 LeuGluLysValGluAlaLaserGlyPheThrSerGluGluLysAspProGluGluPhe 678
Db 1272 CTGGAGAGCGGAGGTGGCCAGGCC--ACAGGCACGTGGGGGAGATAGAGCAGGAG 1328
Qy 679 LeuAsnIleLeu-----PheHisIleLeu 687
Db 1329 CTAGCTCGCCCGGAGCAGCATGACAGCATGTCTCGAATTGGAAGCCAAATGGAC 1388
Qy 688 -----ArgValGluProLeuLeuLysIle 695
Db 1389 CAGCTCGCAACATGTGGAGCTGCTGACAGGAGAGAGGTGGAGCTTCTCAACCCAGCTT 1448
Qy 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPhe 707
Db 1449 GAAGAGGAGAAAGGAAGGTGGAGACCTTCAGTTC 1484
```

RESULT 10

PCT-US05-07748-423

; Sequence 423, Application PC/TUS0507748

; GENERAL INFORMATION:

; APPLICANT: Avalon Pharmaceuticals

; TITLE OF INVENTION: Determining Cancer-Linked Genes and Therapeutic Targets Using

; FILE OF INVENTION: Molecular Cytogenetic Methods

; FILE REFERENCE: 689290-237

; CURRENT APPLICATION NUMBER: PCT/US05/07748

; PRIOR FILING DATE: 2005-03-08

; PRIOR APPLICATION NUMBER: 60/550,304

; NUMBER OF SEQ ID NOS: 3049

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 423

; LENGTH: 5449

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: cdna Sequence

PCT-US05-07748-423

Alignment Scores:

Pred. No.:	13.7	Length:	5449
Score:	146.00	Matches:	112
Percent Similarity:	33.66%	Conservative:	59
Best Local Similarity:	22.05%	Mismatches:	189
Query Match:	2.90%	Indels:	149
DB:	2	Gaps:	21

US-09-671-687A-3 (1-949) x PCT-US05-07748-423 (1-5449)

```
Qy 275 AspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGluSerThr-IleLeuLe 294
|||||
```

```
Db 270 GATGGCAGAGGAGC-----GTGACCAGCACTCACCCCTTGT 305
Qy 294 uHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAl 314
Db 306 CCACCTG-----CCAGTGGCACCCTCAGCAAGCCAGCGGCTGAA 350
Qy 314 aPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaTh 334
Db 351 GCCCCCGCGCGTGGG-----GGGAAGCACTCCACCCCATGGGCGG 392
Qy 334 rGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsnGlySe 354
Db 393 GACATCTACTGGGTGAGCTTCATCTCGCGCGGTGGCGCTAGCTCCAGGAAGGCTC 452
Qy 354 rSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAl 374
Db 453 CCCATGTCACAAACAGTCATCTGGACCTCTCTCTCCCG---GCCGAGCTGTGCTGCC 509
Qy 374 aGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe----- 388
Db 510 CGAAGAGCCGGGCCCCAAGCGCGGGAAGTGGGGGATGACTTCTTGGGGGACTTTTGTGT 569
Qy 389 ---AspArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAs 407
Db 570 GGGCGAGCGGTGTGGGTGAACGCGCTGAAGCCAGCGCTGTGTGAGTATCTGGGAGAC 629
Qy 407 nArgPhe----- 409
Db 630 GCAGTTGCGACCCGGCGCAGTGGGTGGTGTGCTGAGCAGCCGGTGGGCAAGAAATGA 689
Qy 410 -----HisSerLe 412
Db 690 TGGCGCGGTGGCGGTGGCTACTTTCGAGTGCCTCCGCGCTCCAGGGTATCTTCACGCG 749
Qy 412 uProPheSerLeuThrLysMetProAsnThr-----AsnGlySerIleGlyHisSerPr 430
Db 750 GCCTCCAGCTACCCGCGCAGCCAGCGCGGAGGTGCGGGAGTGTATGCCACTCCGT 809
Qy 430 oLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSe 450
Db 810 GGAGTCGCTGACTGCCAGAACCTGTCTCATTCATTGCGGCGACGCGC-----AC 857
Qy 450 rProProLeu-----AlaMetPr 456
Db 858 GCCCCCGCTCACCAGCGCGCTCATCCCTCGCGGAGAGCGCTCTCAACAGCTCCGTGAA 917
Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
Db 918 GACTGCAACAGAGTCGGGATCCAACTCTCAGACAGCGGTCTGTGAAGCGGGGCGAAAA 977
Qy 476 oPheTyr-----GlyValIleArgTr 483
Db 978 GGACCTGGCGCTGGGGAGCCGCTGTGTTGGCGGAGCAGAGACTGGCGTGTGCGGTA 1037
Qy 483 pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCy 503
Db 1038 CGTGGGGAGACAGACTTTGCCAAGGGCGAGTGGTGTGGCGTGGAGCTG---GACGAGCC 1094
Qy 503 sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLy 523
Db 1095 CCTTGGGAAGAATGATGGGCGGTGGCGGGGACCCAGGTACTTCCAGTGCACCACCAAGTT 1154
Qy 523 sAlaLeuPheValLysLeu-----Ly 530
Db 1155 TGGTCTCTTCGCGCCCATCCAAAGTATCCGTATCGGCTTCCCATCTACAGCCACG 1214
Qy 530 sSerCysArgProAspSerArgPheAla-----SerLeuGlnProVa 544
Db 1215 CAAGGCCAAGAAGACCAAGCGTATGGCCATGGGTGTGTGTCAGCACTGACCCACAGTCCAG 1274
Qy 544 lSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGlyGlyTyrLeuSerGl 562
Db 1275 CAGTTTCCTCATAGCTCCGTCAGCTGTGGCTCTCTCGTGGGGGTCTCGGCCAGCGC 1334
```

QY 562 u-----ValValGluGluAsnThrProProLysMetGluLysGluGluLeuGluI 579
 Db 1335 CAGTGGCTGTCTCAGGAGACCTCTTACGCTACGCCCGCAAGATCTCGGACACACGCGC 1394
 QY 579 eMetileGlyLysLysGlyIleGlnGlyHis-TyrAsnSerCysTyrLeuAspSerT 599
 Db 1395 CTTGACGAGGACCTGAAGAGAGACAGACACATTTGAGGAGCTGCTGGCTGAACGAGA 1454
 QY 599 hr-----LeuPheCysLeuPheAlaPheSerSerValLeuAspThr----- 612
 Db 1455 CTTGGAACGGGCTGAG-GTGGCCAAAGGCCACACACATCTCGGAGGTGGAGAGGAGA 1513
 QY 613 --ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGlu- 631
 Db 1514 TTGCCCTGTCTCAAGCACACAGCATGAG-----CAGTATGTTGCAGAAAGCCGAGGAGA 1564
 QY 632 --LeuLeuArgThrGluIleValAlaAsnProLeuArgIleTyrGlyTyrValCysAlaThrL 651
 Db 1565 AGCTGACGAGCGCCGGCTGCTC----- 1587
 QY 651 ysiIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerG 671
 Db 1588 --GTGGAGAGCGTGGGAAA-----GAGAAGTGGACCTGTCCACACAGCTGGAGGAG 1639
 QY 671 luGluLysAspProGluGlu 677
 Db 1640 AGAGGAGGAAGGTGGAGGAT 1659

RESULT 11

PCT-US05-07748-422
 ; Sequence 422, Application PC/TUS0507748
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals
 ; TITLE OF INVENTION: Determining Cancer-Linked Genes and Therapeutic Targets Using
 ; TITLE OF INVENTION: Molecular Cytogenetic Methods
 ; FILE REFERENCE: 689290-237
 ; CURRENT APPLICATION NUMBER: PCT/US05/07748
 ; CURRENT FILING DATE: 2005-03-08
 ; PRIOR APPLICATION NUMBER: 60/550,304
 ; PRIOR FILING DATE: 2004-03-08
 ; NUMBER OF SEQ ID NOS: 3049
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 422
 ; LENGTH: 5554
 ; TYPE: DNA
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: cDNA Sequence
 PCT-US05-07748-422

Alignment Scores:
 Pred. No.: 14 Length: 5554
 Score: 146.00 Matches: 112
 Percent Similarity: 33.66% Conservative: 59
 Best Local Similarity: 22.05% Mismatches: 189
 Query Match: 2.90% Indels: 149
 DB: 2 Gaps: 21

US-09-671-687A-3 (1-949) x PCT-US05-07748-422 (1-5554)

QY 275 AspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGluSerThr-IleLeuLe 294
 Db 270 GATGGCAGAGAGAC-----GTGACACGACCTACCTTGT 305
 QY 294 uHisIleAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAl 314
 Db 306 CCACCTG-----CCGAGTGGCACCGCATGCAGAAAGCCGAGCGGCTGAA 350
 QY 314 aPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaTh 334
 Db 351 GCCCCCCCGCCGCTGGG-----GGGAAGCACTCCAGCCCCCATGGGCCG 392

QY 334 rGlySerThrSerAspProGlyAsnArgSerGluLeuPheTyrThrLeuAsnGlySe 354
 Db 393 GACATCTACTGGTTCAGCTTTCATCTCGCGCGGTGGCGCTAGCTCCAAAGAGGCTC 452
 QY 354 rSerValAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAl 374
 Db 453 CCCATCTGCAACAACAGTCATCTGGACCTCTCTCTCCCGC---GCCGACGCTGCTGCC 509
 QY 374 aGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe----- 388
 Db 510 CGAGAAGCCGGGCCCCCAAGCGCGAGTGGGGGATGACTTCTCGGGGACTTTGTGGT 569
 QY 389 ----AspArgSerSerProProLeuGlnProProProValAsnSerLeuThrThrGluAs 407
 Db 570 GGGCGAGCGGCTGTGGGTGAACGGCGTGAAGCCAGCGCTGTGTGACGTATCTGGGAGAC 629
 QY 407 nArgPhe----- 409
 Db 630 GCAGTTTCGACCGCGCCAGTGGCTGGCTGGTGTGTGGACGACCGGTGGGCAAGAATGA 689
 QY 410 -----HisSerLe 412
 Db 690 TGGCGCGGTGGGCGGTGGCTTTCGAGTCCCGGCCCTCCAGGGTATCTTTCACGCG 749
 QY 412 uProPheSerLeuThrLysMetProAsnThr-----AsnGlySerIleGlyHisSerPr 430
 Db 750 GCCCTCCAAAGCTGACCCCGAGCCCGCGAGGGCTCGGGGAGTGTATCCCACTCCGT 809
 QY 430 oLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSe 450
 Db 810 GGAGTGCCTGACTGCCCAACCTGTCATTGCTGCGGACCGCC-----AC 857
 QY 450 rProProLeu-----AlaMetPr 456
 Db 858 GCCCGCGCTGACCGCGCGGTCTATCCCTCGCGGAGAGCGTCTTCAACAGCTCCGGA 917
 QY 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
 Db 918 GACTGGCAACAGTGGGATCCAACTCTCAGACAGCGGCTCTGTGAAGCGGGCGGCAAAA 977
 QY 476 oPheTyr-----GlyValIleArgTr 483
 Db 978 GGACCTGGCGCTGGGGACCGCGTGTGTGGCGGAGCAAGACTGCGCTGGTGGCGGTA 1037
 QY 483 pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCy 503
 Db 1038 CTTGGGGGAGACAGACTTTCCCAAGGGCGAGTGGTGGCGTGGAGCTG---GACGAGCC 1094
 QY 503 sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLy 523
 Db 1095 CTTTGGGAAGATGATGGGCGGTGGCGGCGACAGGTACTTCCAGTGCACCACCAAGTT 1154
 QY 523 sAlaLeuPheValLysLeu-----Ly 530
 Db 1155 TGGTCTCTTTCGCGCCCATCCAAAGTATCGGTATCGGCTTCCATCTACCACCCGAGC 1214
 QY 530 sSerCysArgProAspSerArgPheAla-----SerLeuGlnProVa 544
 Db 1215 CAAGGCCAAGAGACCAAGCGTATGGCCATGGGTGTGTGACACTGACCCACACATCCAG 1274
 QY 544 lSerAsnGlnIleGluArgCysAsnSerLeuAla-----PheGlyGlyTyrLeuSerGl 562
 Db 1275 CAGTTCTCCATCAGCTCCGTCAGCTCTGTGGCTCTCCGCTGGGGGTGGCGGCGCGC 1334
 QY 562 u-----ValValGluLysLeuThrProProLysMetGluLysGluGlyLeuGluI 579
 Db 1335 CAGTGGCTGTCTCAGGAGACCTTCTTACGCTACGCCCGCAAGATCTCGGCAACACGCGC 1394
 QY 579 eMetIleGlyLysLysGlyIleGlnGlyHis-TyrAsnSerCysTyrLeuAspSerT 599
 Db 1395 CTTGACGAGGACCTGAAGAGAGACGACGACATTTGACGAGCTGTGTGCTGCTGACGAGA 1454
 QY 599 hr-----LeuPheCysLeuPheAlaPheSerSerValLeuAspThr----- 612

-QY 651 yalleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerG 671
Db 1588 --GTGGAGAGCGTGGCGAAA-----GAGAAGGTGGACCTGTCCAACCCAGCTGGAGGAGG 1639
QY 671 luGluLysAspProGluGlu 677
Db 1640 AGAGGAGGAAGGTGGAGGAT 1659
RESULT 13
US-10-932-182A-77607
; Sequence 77607, Application US/109322182A
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77607
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-77607

Alignment Scores:
Pred. No.: 21 Length: 2154
Score: 136.00 Matches: 82
Percent Similarity: 39.38% Conservative: 46
Best Local Similarity: 25.23% Mismatches: 126
Query Match: 2.70% Indels: 72
DB: 7 Gaps: 15

US-09-671-687A-3 (1-949) x US-10-932-182A-77607 (1-2154)

QY 392 SerProLeuGlnProProValAsn-----SerLeuThrThrGluAsnArgPhe 409
Db 412 TCCTTACCGCTACGGCTTTTACCTGTTCATATGCGCTCTTTGCAATACATACAAAGTTT 471
QY 410 -----HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySer 425
Db 472 GATGGCTCACTACATGAATACCAATGAGCTAACAACCAACAATGATAT----- 525
QY 426 IleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAla 445
Db 526 -----TCGAAAGAGGACATAGTCAGGGAAGTAATCAGATA 561
QY 446 ProValGlnSerProProLeuAlaMetProGlyAsnSerHisGlyLeuGluVal 465
Db 562 -----GCATCAAGTAATAAATTCGAAGCT 585
QY 466 GlySerLeuAlaGluValLysGluAsnProPheThrGlyValIleArgTrrIleGly 485
Db 586 GGTTC-----GAAGTG-----GCGTACTACACAGCAAAAGGATTATCA 627
QY 486 GlnProGlyLeu-AsnGluValLeuAlaGlyLeuGluLeuGluAspGluCys----- 503
Db 628 AAACCTTCGTATTAAACTCAGTACTGGAAGAGTCATTATCAAAACGCTGTCTAGT 687
QY 504 -----AlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCy 519
Db 688 CCAGCAACTGCTCCCCCGTCACTCTCTGAGGTTTCGTCAACAATCAGAGATTCTAGC 747
QY 519 sAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAl 539
Db 748 CAAGATTCTCATCTCTTCATCCAGGTGGAACCAACCAAGGAG-GAA----- 795
QY 539 aserLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTy 559

Db 796 -----GAAGGAAGATCGAGGCCATAGATCTTCGCCCCCGAAAGCTTA 839
QY rLeuSerGluValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluI 579
Db 840 CAATCTACCAGTAATTGAAGATCTTAATGATCTGTTATCAGAG-----TTATCTAT 890
QY eMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerTh 599
Db 891 T-----ACTGGCTGCAAAATCCATGCAATACATGCTATATTAATAGTAT 935
QY rLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArgProLysG 619
Db 936 AATTCAATGCTCTTTGGCACTACCTTATTTCGGGATTTCTTTTAAACCAAGAAATATAG 995
QY u-----LysAsnAspValGluTyrSerGluThrGlnGluLeuLeuArgThrGluI 637
Db 996 GCTATTCTCTCAATACTAACAATAATCGAAAGAGTCCAA--CTTCCCGCTCGATT 1052
QY eValAsnProLeuArgIleTyr-----GlyTyrValCysAlaThrLysIleMe 653
Db 1053 CGTTTGTTCAAAAAATGATTGTAATGGAGGAGGAGGATTATACCAATAGATTCTT 1112
QY tLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGlu 673
Db 1113 AAAAATGTGCAAG-----AAGTTAAGGCCAGATTTCGAACATCCAGATGATCAACA 1163
QY sAspProGluGluPheLeuAsnIleLeu-----PheHisIleLeuArgValG 690
Db 1164 GGATACAGGAGTTTCTACTAATAGTTTGTAGCAGCAATACATGAGAGCTCTCGAACA 1223
QY uProLeuLeuLys 694
Db 1224 AATGTTGTCAAG 1236
RESULT 14
US-10-450-763-21942
; Sequence 21942, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hvsseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 21942
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1156)..(995)
; OTHER INFORMATION: 38% homologous to Homo sapiens A secreted protein encoded by
; OTHER INFORMATION: clone dt674_2, accession number W80408, Smith-Waterman Score=87.
US-10-450-763-21942

Alignment Scores:
Pred. No.: 47.4 Length: 1497
Score: 128.00 Matches: 111
Percent Similarity: 34.89% Conservative: 68
Best Local Similarity: 21.64% Mismatches: 183
Query Match: 2.54% Indels: 152
DB: 8 Gaps: 24

US-09-671-687A-3 (1-949) x US-10-450-763-21942 (1-1497)

Db 1863GGGTGTTTCGGCGTTCGCCCAACAGTTTCGGCGTTCGGCAGTTCGGT 1816
Qy 173 GlnGlyPheThrAspGlyValTrpGlnGlyLysGlnLeuPheGlnCysAspGluAspCys 192
Db 1815 GCCATTTTGTGATGGGAGTTTCCCGAAAGATATTCTCAAGCGGCACAGATT 1756
Qy 193 GlyPheValAlaLeuAspLysLeuGluLeuLeuGluAspThrAlaLeuGluSer 212
Db 1755 GAGTTTGTGATGGAAGTCGATGTT----- 1726
Qy 213 AspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProProLeuGluLeuSer 232
Db 1725 -----GATGCGTTTCGATCGCGTCCGAATACTGCCCGCAGCTCTCCAGCGGAATCAC 1675
Qy 233 ArgValSerLeuLys-----GlyGlyGluThr 241
Db 1674 CAGGCTTCGTGATTAAAGCGGTTGTTGGTGTTGGCGATGGCGCGTTCGGCTGCG 1615
Qy 242 IleGlu-----Ser 244
Db 1614 GTCGAGCCAAAAGGTTTTCGGCGCTTCGGCGCATACGGCGATAAAGCTTTCGGCGTCAG 1555
Qy 245 GlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheVal--- 263
Db 1554 GCGCGCGGCAAGTTCGAGATGTTTCGGCGCTGCTCTACGGCGGCTTCGTCGTCGA 1495
Qy 264 -----GlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal 281
Db 1494 AACCAAGTCCGCGCAGCAAGACCATTTTCGTGCTGCTTTCGGCGCGCTTTCGTGGCGTA 1435
Qy 282 LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisAsnAspIleLeuPro 301
Db 1434 GCCGCGGC-----CGCGACATAACGCCA 1411
Qy 302 GluSerValThrGlnGluArgArgPro-----ProLysLeuAlaPhe 315
Db 1410 GTCGAATGCTCCAAACCGCGCGCGCAGCTGTGTCGAGCAGAAATTCGCGGAT 1351
Qy 316 MetSer-----ArgGlyValGlyAspLysGlySerSerHisAsnLysProLys 332
Db 1350 TTCGAGGATAGAGGCGTGGCGCGCGCGCTGCGCGGCGGCGGCGGCGGCGGCGGCGG 1297
Qy 333 AlaThrGlySer-----ThrSerAspProGlyAsnArg 343
Db 1296 -----GGTGGCGGTGATTTCGGCATTTTGTCAATACGAGGCAACCGCGTAATGAT 1243
Qy 344 ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLys 363
Db 1242 GCCGTCGCTACCTTCTTCTGACGCGCGCGCGAGTTCGCT-----CAAAAATTT 1195
Qy 364 SerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGlu 383
Db 1194 GTCGGTAACGCTTTCGCCAAACCGGACTTTCGGAA-----TTT 1156
Qy 384 IleSerThrAspPheAsp-ArgSerSerProProLeuGlnProProValAsnSerLe 403
Db 1155 GTCTCCGGGATTTCCAGCGTTTCGGTTTAAACGATGTTGATCCGCTCTCAATCCAGCGT 1096
Qy 403 uThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAs 423
Db 1095 GTGACGTCGAAACACGCGCATTTCTTCTGCTGTTGATTTTTCGCGAAATTTGTCGCGCACGCG 1036
Qy 423 nGlySerIleGlyHisSerProLeuSerLeuSer-----AlaGlu 436
Db 1035 TTCGATACGAGCATTCGCTTTCGAGGTTTAAACGATGTTCCAGTAGCGAGGTTGTCCAA 976
Qy 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456
Db 975 GGCAGTACCCAC-----AATACGCTTTTTCACCGCGCATTCATCGCCACATTCGC 922
Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476

Db 921 GCCACGCGAGGACGCGTTCGGCGGAAAGTTCGATCGACGCGC----- 882
Qy 476 oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlu 496
Db 882 ----- 882
Qy 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516
Db 881 -----AACACCAAGCC 871
Qy 516 xPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536
Db 870 CCGCTGATCGCGGTTTCTTCGACCGCGCGGTAAACACCGCGCGCGC-CCACCGGATAA 812
Qy 536 xArgPheAlaSerLeuGln-ProValSerAsnGlnIleGluArgCysAsnSerLeuAlaP 556
Db 811 TCGGATGTCGCGTACAAGCTTCGCGC----- 783
Qy 556 heGlyGlyTyrLeuSerGluValGluAsnThrProProLysMetGluLysGluG 576
Db 783 ----- 783
Qy 576 yLeuGluLeuMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrL 596
Db 782 -----TCAACGTAT- 774
Qy 596 euAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeu-Leu 615
Db 773 -----TCCACGCGCGATGCTTGTCCAGCTTT--TCCGTATTGATGATCGCTGTTG 722
Qy 616 ArgProLys-----GluLysAsnAs 622
Db 721 CGTCCAAAGGTACTGCGCGCGCGGTATAACCGCTACCGCGCGCGCGGATAATAACCA 662
Qy 622 pValGluTyr-----TyrSerGluThrGlnGluLeuLeuArgThrGluIleVa 638
Db 661 AGTCCAAATTCGATTAGGCGCGCGCCCAAGCGCGACTTCGCGCTCCGTTCGGGATGA 602
Qy 638 lAsnProLeuArgIleTyr-----GlyTyrValCysAlaThrLy 651
Db 601 CGACGACAAACGGATATTACGCGCGCAGTCGCGTAAACGTGTGTACCOCG 542
Qy 651 s-----IleMetLysLeuArgLysIleLeuGluLysValGluAlaAl 665
Db 541 CCAGCCCGTCAACATAATATTGCGCGTGTGTAATTTTCTCAACCGCTCCAA- 487
Qy 665 aserGlyPheThrSerGluLysAspProGluGluPheLeuAsnIleLeuPheHisHi 685
Db 487 ----- 487
Qy 685 sIleLeuArgValGluProLeuLeu-----LysIleArgSerAlaGlyGlnLysValGlnAs 704
Db 486 -----AATCTCGCGCGCTTTTGGCGGTTTCGTCAAAGCTCCGCTCAAAACGCTCCA 434
Qy 704 pCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlu 724
Db 433 CTGC-----CTTTTCGCGCGCGCAATCAACACATCGCTTGTGATATCGT 386
Qy 724 nGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCy 744
Db 385 CCGCGCGTTCGCG-----GATTTCTGTTCAACCTGTGGCGCATTTTCGCGTACCAACCGCGCGC 329
Qy 744 s-----LeuIleIleGlnMet-----ProAr 751
Db 328 GCGGTTTCGATGCTCCAGCAATCATCGACAGATACGGATTTCGCGCACGACCCCAA 269
Qy 751 gPheGlyLysAspPheLysLeuPheLysIlePheProSerLeuLeuAsnIleTh 771
Db 268 TATCGGCCAACACTTCAACACACATCGCGCGCAACCGCGGTTTTCGCTGTCC---GC 212
Qy 771 rAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGly----- 786
Db 211 GCAAGTCTTCGCAATGTGCCACGCTCGT-CGCCCCAGCAGCGGATGACGANTTCGCG 153

```
QY 787 -----LeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGly-- 804
Db 152 TCGGTATAAGAAAGTATAGTTGTAGGGAATTT-----CCGAAATACGCTCGGGGCG 102
QY 805 -----LysIleLysGlnPheCysLysThrCysAsnThrGlnVa 817
Db 101 GTAGTCGTGTCATGTCGTGTCGGTGTTCGGCAGTTTTTTATTGAAATGC----- 53
QY 817 lHisLeuHisProLysArgLeuAsnHisLysTyrAsnPro 830
Db 52 -----CGTTTGAAAAGTCGGTTTCAGCCG 29
```

Search completed: April 18, 2005, 22:23:19
Job time : 650 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:21:14 ; Search time 56 Seconds
(without alignments)
1265.035 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 5034
Sequence: 1 MSSGLWSQEKVTSPYWEERI.....RLLCDAYMCVQSPWLSLYK 949

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A COMB pep.*
2: /cgn2_6/ptodata/1/1aa/5B COMB pep.*
3: /cgn2_6/ptodata/1/1aa/6A COMB pep.*
4: /cgn2_6/ptodata/1/1aa/6B COMB pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS COMB pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	3.0	1427	US-09-538-092-1044	Sequence 1044, Ap
2	135	2.7	1043	US-09-949-016-11650	Sequence 11650, A
3	130.5	2.6	484	US-09-248-796A-24175	Sequence 24175, A
4	125.5	2.5	1469	US-09-262-537-58	Sequence 58, Appl
5	118.5	2.4	1466	US-09-262-537-20	Sequence 20, Appl
6	118.5	2.4	1471	US-08-811-519-1	Sequence 1, Appli
7	118	2.3	267	US-09-399-913-57	Sequence 57, Appl
8	118	2.3	267	US-09-350-614-57	Sequence 57, Appl
9	118	2.3	677	US-09-270-767-45699	Sequence 45699, A
10	117	2.3	386	US-09-248-796A-17565	Sequence 17565, A
11	117	2.3	798	US-08-222-617A-8	Sequence 8, Appli
12	115.5	2.3	719	US-09-949-016-7766	Sequence 7766, Ap
13	114	2.3	523	US-09-248-796A-17555	Sequence 17555, A
14	114	2.3	850	US-09-129-603-2	Sequence 2, Appli
15	114	2.3	2777	US-10-220-587-4	Sequence 4, Appli
16	113	2.2	1226	US-08-280-443-2	Sequence 2, Appli
17	113	2.2	1226	US-08-457-459-2	Sequence 2, Appli
18	113	2.2	1226	US-08-555-678-2	Sequence 2, Appli
19	113	2.2	1226	PCT-US95-02275-2	Sequence 2, Appli
20	112.5	2.2	493	US-08-999-774A-12	Sequence 12, Appl
21	112.5	2.2	805	US-09-538-092-257	Sequence 257, App
22	112	2.2	734	US-09-328-352-4412	Sequence 4412, Ap
23	112	2.2	869	US-09-902-540-10125	Sequence 10125, A
24	111.5	2.2	1317	US-09-949-016-7588	Sequence 7588, Ap
25	110.5	2.2	657	US-09-370-368-7	Sequence 7, Appli
26	110	2.2	577	US-09-949-016-11572	Sequence 11572, A
27	109.5	2.2	868	US-09-538-092-787	Sequence 787, App

28	109.5	2.2	1306	3	US-08-999-774A-13	Sequence 13, Appl
29	109	2.2	452	4	US-09-205-258-689	Sequence 689, App
30	109	2.2	667	4	US-09-949-016-7759	Sequence 7759, Ap
31	108.5	2.2	493	4	US-09-538-092-1210	Sequence 1210, Ap
32	108	2.1	533	4	US-08-216-592A-4	Sequence 4, Appli
33	108	2.1	665	4	US-09-328-352-6983	Sequence 6983, Ap
34	108	2.1	753	4	US-09-949-016-6676	Sequence 6676, Ap
35	107.5	2.1	659	4	US-09-562-737-18	Sequence 18, Appl
36	107.5	2.1	897	4	US-09-538-092-315	Sequence 315, App
37	107.5	2.1	1477	4	US-09-206-942-71	Sequence 71, Appl
38	107.5	2.1	2616	6	5206163-3	Patent No. 5206163
39	107.5	2.1	2616	6	5206163-3	Patent No. 5206163
40	107	2.1	501	4	US-09-640-211A-1027	Sequence 1027, Ap
41	106.5	2.1	696	3	US-07-757-342D-4	Sequence 4, Appli
42	106.5	2.1	696	4	US-09-461-657B-4	Sequence 4, Appli
43	106.5	2.1	806	4	US-09-949-016-7572	Sequence 7572, A
44	106.5	2.1	1049	4	US-09-248-796A-18611	Sequence 18611, A
45	105.5	2.1	491	4	US-09-302-626B-191	Sequence 191, App

ALIGNMENTS

RESULT 1

US-09-538-092-1044
; Sequence 1044, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 1044
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P30622
US-09-538-092-1044

Query Match	3.0%;	Score 151;	DB 4;	Length 1427;
Best Local Similarity	20.5%;	Pred. No. 0.00026;		
Matches	99;	Conservative 55;	Mismatches 171;	Indels 157; Gaps 20;
QY	260	GYFVGVDMPNIGNWDGRFDGVLCSFACVSESTILLHINDIIPESVTOERRPPKLAFMSRG	319	
Db	85	GQWAGVLDFEPIKNDGVSAGVR-YFQCEP-----LKGIF-----TRPSKL---TRK	127	
QY	320	VGDKGSSSHNKPKATGSTDGPNRRSELFTYLGSSVDSQPSQSKNTWYDEVAEDPAK	379	
Db	128	V-----QAEDEANGLQTTFASRATSPCLTSTASMVSSTPSFN-----IPQKPSQPA-	175	
QY	380	SLTEISTDFDRSSPPPLQPPPVNSITTENRRFSLPFLSTKMTNTNGSIHGSPLSLSAQSV	439	
Db	176	-----AKEPSATPPISNLT-----KTASESIS	197	
QY	440	EELNTPAVQESPPPLAMPNGNSHGLEVGSLEAVKENPPFYGVIRWIGOPPGNEVLAGLE	499	
Db	198	NLSEAGSIKK-----GERELKIGDRVLVGTKA--GWRFLETDFAKGWCWGVEL	246	
QY	500	EDCAGCTGDTFTGTYFTTCAKALFVKULKSCRPSDRFASLQPVSNQIERCNLSLAFGGY	559	
Db	247	-DEPLGKNDGAVAGTRYFQCPKYGLFA-----PVHKVTKIGF----	283	

QY 560 LSEVVEENTPPKMEKGL-EIMGKKGKQGHVNSCYLDTLCLFAPSSVLDTVLLRPK 618
Db 284 -----PSTTPAKAKANAVRRVMAATTSASLKRSPSASSLS-----MSSVASSVSRP- 330
QY 619 EKNDVEYSETQELLATEINWPIRIYGVYCATIMK-----LRKILEVEAASGF 668
Db 331 -----SRTGLTTSRYARKISGTTALQALKEKQOHIEQLLAERDLERAFAKA- 381
QY 669 TSEKPEEFLNLL-----FHHIL-----RVEPLLKIRSAQKQVQDC 705
Db 382 TSHVGEIEQALARDGHDQHVLEAKMDQLRTWVEADREKVELLNQLEBEKRVEDL 441
QY 706 YF 707
Db 442 QF 443

RESULT 2

US-09-949-016-11650
; Sequence 11650, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11650
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11650

Query Match 2.7%; Score 135; DB 4; Length 1043;
Best Local Similarity 22.6%; Pred. No. 0.0051;
Matches 66; Conservative 26; Mismatches 112; Indels 88; Gaps 10;
QY 309 RPKLAFMSRGVGDGKSSSHNKPATGSTDGPNRRSELFTYTLNGSSVDSQPSKNTW 368
Db 8 KPP-----GRGK-----HSSPMGRSTSGASSAAVAASKEGSPHLKQSGSPSSP- 55
QY 369 YIDEVAEDPAKSLTEISTDF-----DRSSPPLQPPVNSLTENRF----- 409
Db 56 AAAAAPEKPGKAAEVDGDFLGVGERVYVGVVQYIGETQFAPCQWAGVVLD 115
QY 410 -----HSLPFLSLTKMPT--NGSIGHSPLSLSAQSVMEELNT 444
Db 116 PVGNKDGAVGVRYFECPALQGIETREPSKLTROPTAEGSGSDAHSVESLTQNLHSGT 175
QY 445 APVQESPL-----AMPPGNSHGLEVGSLAEVKENPPFY----- 478
Db 176 A-----TPPLTSRVTLRESVLNSVKTGNSGSLSDSGSVKRGKDLRLGDRVLVGGTK 231
QY 479 -GVIRWIGQPPGNEVLVAGLEDEACAGCTDGTFRGTRYFTCAKXKALFKVL 529
Db 232 TGVVRYVGETDFAKGECGVLE-DEPLGKNDGAVAGTRYFCPPKFLGFAPI 282

RESULT 3

US-09-248-796A-24175
; Sequence 24175, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24175
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24175

Query Match 2.6%; Score 130.5; DB 4; Length 484;

Best Local Similarity 22.2%; Pred. No. 0.0035;
Matches 128; Conservative 68; Mismatches 187; Indels 193; Gaps 35;

QY 317 SRGVGDKGSSHNKPKATG-----STSDPGNRRSELFTYTLNGSS-----VDSQPSKS 364
Db 39 SSNLDNSGTT-----KTSGHEIVNISKVPNGNEHSE--YEVNGVGLSILEHSSPFPKR 92
QY 365 KNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTENRFHSLP--FSLTKMPT 422
Db 93 KTTYQERVSYIPNHDFTGKSSIFESSP--QP-----SIPNTFYKSHRNA 137
QY 423 NG-SIGHSPLSLSAQSVMEELNTAPVQESPLA--MPPGNSHGLEVGSLAEVKENPPFYG 479
Db 138 KAKALGSTQITEETSQEQEQN-GDVQ--PPLSTKQTPKNN-----KENT----- 180
QY 480 VIRWIGQPPGNEVLVAGLEDEACAGCTDGTFRGTRYFTCAKXKALFKVLSKCRPDSRFA 539
Db 181 TPRWM--PPVLDEKMNVDLED-----IKFT-----PSMPSIQPHH-- 215
QY 540 SLOPVSNQ-----IERCN--SLAFGGY--LSEVVEENTPPKMEK 574
Db 216 -QQPLSNQSRGGISPSPFQDFEHTGDTMIHNSVQSVETPGYRRARQDYKQKQPDMLK 274
QY 575 EGLEIMIGKKGIQG-----HNSCYLDLSTLCLFAPSSVLDTVLLRPKE-KND 622
Db 275 SIPTVGGSGSGGGAGPDETNTKDHSTS-----STI-----SSMFGTIQKRPEEIKQ 324
QY 623 VEYVSETOELLRTVINPLRIYG--YVCATKIMKILKEKVEAASGFTSEKDPPEFLN 680
Db 325 METINEMKMNKPE--SPLKLYGPKYNTYTR-NQLAGLVENLKNKNTPAQNQONQ--- 378
QY 681 ILFHHILRVBPLLKIRSAQKQVQDQYFYQIFMEKNEKVGVTIQQLL-EWSFINSNLKFA 739
Db 379 -----NQQN-----NQNQATSTSPNSVLSKPPALNLSKIV 412
QY 740 EAPSCILIQMPREGK-DFKLFK-----IPFSLNELITDLEDTPRCRIQGLAMVSECRE 794
Db 413 NTPPKNIKSTTKTGAYDEKSYLNAENIFFILK--VKDLRSIT-TTTTIIMVLSLYHRR- 468
QY 795 CYDDPDISAGIKQFCCKTQNTQVHLHPKRLNHNKYNP 830
Db 469 -----QTHLLPRRKLVMVNP 483

RESULT 4

US-09-262-537-58
; Sequence 58, Application US/09262537
; Patent No. 6479256
; GENERAL INFORMATION:
; APPLICANT: Havlick, Joel
; TITLE OF INVENTION: Lectomedin Materials and Methods
; FILE REFERENCE: 27866/35307
; CURRENT APPLICATION NUMBER: US/09/262,537
; CURRENT FILING DATE: 1999-03-04
; EARLIER APPLICATION NUMBER: 60/076,782

EARLIER FILING DATE: 1998-03-04
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 58
LENGTH: 1469
TYPE: PRT
ORGANISM: Homo sapiens
US-09-262-537-58

Query Match 2.5%; Score 125.5; DB 4; Length 1469;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 130; Conservative 80; Mismatches 220; Indels 199; Gaps 35;
QY 172 GGGFTDGVYQKQLFCQDCGFCVALDKLELIEDDDTALESYAGGDTMQLVPLEIN 231
DB 208 GTGFFV--VYDGAVALYKERTRNIVKYD-----WGGKTDLDAVDENGLWV 232
QY 232 SRVSLKGGGTIESGTIVFCQDVLPGKESLGVFGVDMNDPIGNWDGRFD-----GVLC 283
DB 233 LTRIKSGETV-INTANYHDTSPYR-----WGGKTDLDAVDENGLWV 274
QY 284 SPACVESTILLHINDIIPESVTQE-----RRPKLAFMSRG-----VGDKGSS 326
DB 275 IYATEGNGRLVVSQNLNPTLRFEGTWTGYDKRSASNAFMVCGVLYLRSVYVDDSEA 334
QY 327 SHNPKATGSTDPGNRRSELFTYTLNG-----SSVDSQPSKSKNTW--YIDEVAEDPAK 379
DB 335 AGNRVDYAFTN--ANREEPVSLTFPNPYQFISVVDYNDPRDNLQYVNNYF-----383
QY 380 SLTETSTDP--DRSSPPLQPPVNSLTENRHSPLFSLTKMPNTNGSIHSPLSLSAQ 436
DB 384 -VVRYSLFEGPPDPSPAGPATSPPLSTTTTA---RPTPLTSTASPAATTLRRAPLTHPV 439
QY 437 SYMELN-----TAPVQSPPLAMPNGSHGLEVGSLAEVKENPPFYGVIRWIGOPPGL 490
DB 440 GAINQLGPLPATADV---PSTRPPAPNLHVSPLFCEPRE-----VRRVQWP---486
QY 491 NEVLAGELEDECAGCTDGTGTRGYFTCALKKALF---VKLKSCRDPDSRFASLQPVSN 546
DB 487 -ATQGMGLVE---RCPKGT-RGIASFQCLPALGLWNPGRGPDLSNC-----TSPWVN 533
QY 547 QI-----ERCNSLA-----FGYLSEVVEENTPPKMEKEGLEIMIGKKGIQ 588
DB 534 QVAQKIKSGENANIASELARHTRGSIYAGDVSSV-----KLMQLLDILDAQLALRP 587
QY 589 -----GHYNSCYLDSTLFCIFAFSSVLDTV--LLRPKEKNDVEYVYSETQELLRTEIV 638
DB 588 PIERESAGKNYNMVKRERT-CKDYIKAVVETVDNLRLP-----EALSWKDMNATEQV 640
QY 639 NPLRIYGVVCAKIMKRLKILEKVEAASGFTSEKDPPEFLNLFHHLRVEPLLKIRSA 698
DB 641 H-----TATMLL---DVLE--EGAFLLADNVREPARFLAAKENVVLEVTVL---NTE 684
QY 699 GQVQDCVYQO-IFMEKNE-KVGVPITIQ 725
DB 685 GQ-VQELVFPQBEYPRKNSIQLSAKTIQ 712

RESULT 5

US-09-262-537-20
Sequence 20, Application US/09262537
Patent No. 6479256
GENERAL INFORMATION:
APPLICANT: Haylick, Joel
TITLE OF INVENTION: Lectomedin Materials and Methods
FILE REFERENCE: 27866/35307
CURRENT APPLICATION NUMBER: US/09/262,537
CURRENT FILING DATE: 1999-03-04
EARLIER APPLICATION NUMBER: 60/076,782
EARLIER FILING DATE: 1998-03-04
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20

LENGTH: 1466
TYPE: PRT
ORGANISM: Rattus rattus
US-09-262-537-20

Query Match 2.4%; Score 118.5; DB 4; Length 1466;
Best Local Similarity 20.6%; Pred. No. 0.35;
Matches 117; Conservative 80; Mismatches 207; Indels 163; Gaps 31;
QY 233 RVSLKGGGTIESGTIVFCQDVLPGKESLGVFGVDMNDPIGNWDGRFD-----GVLC 284
DB 234 RTRIKSGETV-INTANYHDTSPYR-----WGGKTDLDAVDENGLWV 275
QY 285 PACVESTILLHINDIIPESVTQE-----RRPKLAFMSRG-----VGDKGSS 327
DB 276 YATEGNGRLVVSQNLNPTLRFEGTWTGYDKRSASNAFMVCGVLYLRSVYVDDSEA 335
QY 328 HNKPKATGSTDPGNRRSELFTYTLNG-----SSVDSQPSKSKNTW--YIDEVAEDPAK 380
DB 336 GNRVDYAFTN--ANREEPVSLTFPNPYQFISVVDYNDPRDNLQYVNNYF-----383
QY 381 LTRISTDF--DRSSPPLQPPVNSLTENRHSPLFSLTKMPNTNGSIHSPLSLSAQ 437
DB 384 VVRYSLFEGPPDPSPAGPATSPPLSTTTTA---RPTPLTSTASPAATTLRRAPLTHPV 440
QY 438 VMEELN-----TAPVQSPPLAMPNGSHGLEVGSLAEVKENPPFYGVIRWIGOPPGLN 491
DB 441 AINQLGPLPATAP--APSTRPPAPNLHVSPLFCEPRE-----VRRVQWP---486
QY 492 EVLAGLELEDECAGCTDGTGTRGYFTCALKKALF---VKLKSCRDPDSRFASLQPVSNQ 547
DB 487 ATQGMGLVE---RCPKGT-RGIASFQCLPALGLWNPGRGPDLSNC-----TSPWVNQ 534
QY 548 I-----ERCNSLA-----FGYLSEVVEENTPPKMEKEGLEIMIGKKGIQ- 588
DB 535 VAOIKIKSGENANIASELARHTRGSIYAGDVSSV-----KLMQLLDILDAQLALRP 588
QY 589 -----GHYNSCYLDSTLFCIFAFSSVLDTV--LLRPKEKNDVEYVYSETQELLRTEIV 639
DB 589 IERESAGKNYNMVKRERT-CKDYIKAVVETVDNLRLP-----EALSWKDMNATEQV 641
QY 640 PLRIYGVVCAKIMKRLKILEKVEAASGFTSEKDPPEFLNLFHHLRVEPLLKIRSA 699
DB 642 -----TATMLL---DVLE--EGAFLLADNVREPARFLAAKENVVLEVTVL---STE 684
QY 700 QKQDCVYQOIF-MEKNEKVGVPITIQ 725
DB 685 GQVQELVFPQBEYASESIQLSANTIKQ 711

RESULT 6

US-08-811-519-1
Sequence 1, Application US/08811519B
Patent No. 6630345
GENERAL INFORMATION:
APPLICANT: Petrenko, Alexandre
TITLE OF INVENTION: CALCIUM INDEPENDENT RECEPTOR OF ALPHA-LATROTOXIN,
CHARACTERIZATION AND USES THEREOF
FILE REFERENCE: 1049-1-007
CURRENT APPLICATION NUMBER: US/08/811,519B
CURRENT FILING DATE: 1997-03-04
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1471
TYPE: PRT
ORGANISM: rat
US-08-811-519-1

Query Match 2.4%; Score 118.5; DB 4; Length 1471;
Best Local Similarity 20.6%; Pred. No. 0.35;
Matches 117; Conservative 80; Mismatches 207; Indels 163; Gaps 31;

[illegible]

	Query Match	2.3%;	Score 118;	DB 3;	Length 267;
	Best Local Similarity	34.3%;	Pred. No. 0.019;		
	Matches	46;	Conservative 16;	Mismatches 54;	Indels 18; Gaps 8;
Qy	217	PGDTMQVELPPLRNSVSLKGGTETTESGVIFCDVLPGKESLYGFVGVDMDNPIGNWDG	276		
		:	:	:	:
Dd	104	PGNIMLSAL-GLRLGRVLLDGOKT---GTLRFCGTT--EFASGQWVGVELDPEGNKDQ	157		
		:	:	:	:
Qy	277	RFDGV---LCS-----FACVESYILLHINDIIPESVTQERRPPKLAFMSRGVGDKGSSSH	328		
		:	:	:	:
Dd	158	SVGGVRYFICPKQGLFASYSK--VSKAVDAPSSSVTSTPTPRMDF-SRVTG-KGRREH	213		
		:	:	:	:
Qy	329	NKPATGSTSDPGN	342		
		:	:	:	:
Dd	214	KGKKSPSPSLGS	227		
		:	:	:	:

```

RESULT 8
US-09-350-614-57
; Sequence 57, Application US/09350614
; Patent No. 6689581
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Huai-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: WNI-070CP
; CURRENT APPLICATION NUMBER: US/09/350,614
; CURRENT FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-350-614-57

```

Query Match	2.3%;	Score 118;	DB 4;	Length 267;
Best Local Similarity	34.3%;	Pred. No. 0.019;		
Matches 46;	Conservative 16;	Mismatches 54;	Indels 18;	Gaps 8;
Qy	217	PGDTMQLVLEPLLEINSRVSLKGGTETIESGVIFCDVLPGKESLGYFVGVDMDNFIQNWGD	276	
Db	104	PGNLMLSAL-GLRLGDRVLIDGQKT-----GTURFCGTT--EFASQWVGVELDEPEGKNDG	157	
Qy	277	RFDGV---LCS-----FACVESTILLHINDIIPESVTQERRPPKLAFLMRSGVGDKGSSSH	328	
Db	158	SVGGVRVFI CPPKQGLFASVSK--VSKAVDAPSSSVTSTPTPRMDF-SRVTG-KGRRH	213	
Qy	329	NKPKATGSTDPGN	342	
Db	214	KGKKKSPSSPSLGS	227	

RESULT 9
US-09-270-767-45699
; Sequence 45699, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767

Query Match 2.3%; Score 118; DB 4; Length 677;
Best Local Similarity 20.5%; Pred. No. 0.1;
Matches 81; Conservative 59; Mismatches 154; Indels 102; Gaps 18;
QY 196 ALDKLEIEDDDTALESYAGPGDTMQLVPLP--EINSRVSLKGGTIESGTVIFCDV 255
DB 226 ALESITMCDGTLIKSHIGQIDEISLTSK-----TASGCTMDNST----- 269
QY 256 KESLGYFVGVDMDNPIGNWDGRFDGVLCSFAC--VESTILL-----HINDIIPES 303
DB 270 -DSMSIPLAADRTMP-----VVSAVCPALQRTMILGEMIGDITFNLVDSLTS 318
QY 304 VTQERRPPKLAFWSRGVGDKSSSHNKPATGTSDPGNRRSELFVTLGSSVDSQPQSK 363
DB 319 ALQSE-----SRSLPDGNATFKRPTASATADETQ-----VLTG-----R 354
QY 364 SKNTWYIDEVAEDPAKSLTEISDFDRSSPPQLQPPVNSLTENRPHSLPFSLTQMPNTN 423
DB 355 QMNTFTDGC--NTPGRCETENIDRKLALLTMESSTPLTNNRSHCVHNN-----NNN 407
QY 424 GSIGHSPLSLSAQSVMEELNAPVQSPPLAMPNGSHGLEVSLEAVKENPPFYGVIRW 483
DB 408 NKAGYTP--TLKRG-----DMNLSPVIGATP-QKPTGTAPGRLANTFEPVAKTAPFNGE--- 459
QY 484 IQQPPGLNEVLAGELEDECAGCTDCTF-----RGTRYFTCALKKALFVKLKS CRDSRF 538
DB 460 -----KFLVDTMELLEQIEQLDGTYNLQWSBQHRQMCVNDLA-----BAEVEW 504
QY 539 ASLQPVSNQIERCNSLAFGGYLSVEVVENTPPKMEK 574
DB 505 LAQQDEEQFE--NMLAELGKNTLNEEQI--KMQK 536

RESULT 10
US-09-248-796A-17565
; Sequence 17565, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17565
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (21)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknd
US-09-248-796A-17565

Query Match 2.3%; Score 117; DB 4; Length 386;
Best Local Similarity 18.8%; Pred. No. 0.046;
Matches 75; Conservative 63; Mismatches 162; Indels 100; Gaps 17;
QY 194 FVALDKLEIEDDDTALESYAGPGDTMQLVPLP--EINSRVSLKGGTIESGTVIFCDV 252
DB 2 FTTPEIQIMSSSSSSSSXTTATTTTSARILRPSISELTSRSTISGGSSNNGSALKSQI 61
QY 253 LPCKESLGYFVGVDMDNPIGNWDGRFDGVLCSFACVESTILLHINDIIPESVTOERRPPK 312
DB 62 SPRLSDTSRILPSILKNTSGS-----ST-----PTSSSTPFKCPP 96
QY 313 LAFMRGVGDKSSSHNKPAT--GSTSDPGNRRSELFVTLGSSVDSQPQSKNTWYID 371
DB 97 I-----KSTVGGTSSGNTQSN--VVLGNTKINSPLRSLSSPTLPKAV 136
QY 372 EVAED-----PAKSLTEISTDFDRSSPPQLQPPVNSLTEN-----RHSLPF----- 414
DB 137 QPOCQPHLPASSLSFVTRVI--NTPPQPSQSVASSTSPNTQYQYVYQQQSSPTQQQQQ 194
QY 415 -----SLTKMPNTNGSIGH--SPLSLSAQSVMEELNAPVQSPPLAMPNGSHGLEV 465
DB 195 QQQATPAATTVNMQAQNQSPHAPLQYATQYYPQ-----PVYQSPAGVPPPPSVTHQ 250
QY 466 GSLEAVKENPPFYGVIRWIGQPPGLNEVLAGELE-----EDEACGCTDCTF-----R 512
DB 251 GHIIAVHQHGHLPQGVNGMPPNVGYTIQPEIVNKSTNRCHRC--GTTTPEWRRGPK 308
QY 513 GTRYF--TCALKKALFVKLKS CRPDSRFASLQPVSNQIER 550
DB 309 GVRTCLNACGLFHAFLVK-----RKGALAAAEVILNKKVTK 344

RESULT 11
US-08-222-617A-8
; Sequence 8, Application US/08222617A
; Patent No. 5882879
; GENERAL INFORMATION:
; APPLICANT: Veenstra, Annemarie B.
; APPLICANT: Martin, Juan F.
; APPLICANT: Garcia, Bruno D.
; APPLICANT: Gutierrez, Santiago
; APPLICANT: Barredo, Jose L.
; APPLICANT: Von Doehren, Hans
; APPLICANT: Palissa, Harriet
; APPLICANT: Van Liempt, Henk
; APPLICANT: Montenegro, Eduardo P.
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
; TITLE OF INVENTION: Quantities of ACV Synthetase
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,617A
; FILING DATE: 04-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 97,157
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

Db 1748 SANITVEILPDEDELDKAFSVSVSSGSLGAHINAILTVLASDDPYGIFIFPEKNRP 1807
QY 271 IG-----NWDGRFDGVLCSFACVESTILLHINDIIPESVTQERR---PPKLA 314
Db 1808 VKVEEATQNTLSITLKGMLMGKVLVSATLDA-----MEKPPYFPFNLA 1852
QY 315 FMSRG---VGDKGSS--SHNKRKATGSTS---DPGNRRSELYTLNGSSVDSQPQSKS- 364
Db 1853 RATQGRDYIPASGFALFGANQSEATIAISILDDDEPERSESFVIELNGLTLVAKVQSRSI 1912
QY 365 -----KNTWYIDEVAEDPAKSLTEISDTDFRSSPPLQPPPVNSLTNRRPHSLPF 414
Db 1913 PNSPLRGPVKVETIAQLIIIIANDDAFGTLQLSAPIVRAENHVGPPIINVTRTGGAFADSV 1972
QY 415 SLTKMPTNGSIGHSPLSLSAQSVM-----EELNTPVQESPPPLAMPNGNSH----- 461
Db 1973 KFKAVPIT--ATAGEDYSIASSGVVLEGETSKAVPIYVINDIYPBLGESFLGQLMNETT 2030
QY 462 -GLEVGS LAE-----VKENPPFYGV-----I 481
Db 2031 CGARLGALTEAVIIIEASDDPYGLFGFQITKLIVERPEFNSVKVNLPIIRNSGTLGNVTV 2090
QY 482 RWI-----GO-----PFG-----LNEVLG-----LELEDECAGC 506
Db 2091 QNVATINGQLATGDLRWVSGNVTFAFGETITQLLLEVLADDVPEIBEVIQVQLTDASGGG 2150
QY 507 TDGTFRGTRYFTCALWKALFVKLSCRPDSRFPASLOPVSNQIERCNSLAFGGYLSEVVEE 566
Db 2151 TIG-----LDRIANIIIPANDDPYGTVAFAQVY 2179
QY 567 NTPPKMEKEGLEIMIGKKKGIOGHYNSCYLDSTLFCFLFAPSSVLDTVLLRPKEKNDEVEY 626
Db 2180 RVQEPLEERSYANITVRRSG--GHFGR-----LLLFYSTSDIDVVALAMEEGQDLLSY 2230
QY 627 SET-----QELLRT-----EIVNPLRIYGVCA TKMKILEKVEAASGTSEEKDP 675
Db 2231 YESPIQGVDPDLWRVTWNVSAVGEPL-----YTCATILCLK-----EQACSAFSPFSASEGP 2281
QY 676 EEF 678
Db 2282 QRF 2284

Search completed: April 18, 2005, 15:26:33
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:33:50 ; Search time 376 Seconds
(without alignments)
4129.858 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 5034
Sequence: 1 MSSGLWSQKVTSPWEERI.....RLLCDAYMCYQSPWLSLYK 949

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DB=us-09-671-687a-3 -DEV=us-09-671-687a-3

-MODEL=frame+ p2n.model -DEV=us-09-671-687a-3
-Q=/cgn2_1/USPTO_spool_p/US09671687/runat_18042005_115019_17155/app_query.fasta_1.1095
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687 @CGN 1 1 105 @runat_18042005_115019_17155 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUTUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3676	73.0	2523	4 US-09-620-312D-290	Sequence 290, Appl
2	2468	49.0	2116	4 US-09-646-403-3	Sequence 3, Appl
3	185	3.7	110	4 US-09-513-999C-20085	Sequence 20085, A
4	151	3.0	5857	4 US-09-220-132-79	Sequence 79, Appl
5	139	2.8	2825	4 US-09-949-016-1895	Sequence 1895, Ap
6	139	2.8	3130	4 US-09-949-016-5779	Sequence 5779, Ap
7	135	2.7	2643	3 US-09-399-913-56	Sequence 56, Appl
8	135	2.7	2643	4 US-09-350-614-56	Sequence 56, Appl
9	133	2.6	3245	4 US-09-774-528-352	Sequence 352, App
10	130.5	2.6	1452	4 US-09-248-796A-10072	Sequence 10072, A
11	128.5	2.6	5610	4 US-09-262-537-57	Sequence 57, Appl
12	126	2.5	1503	3 US-08-999-774A-11	Sequence 11, Appl

13	126	2.5	4282	4 US-09-976-594-799	Sequence 799, Appl
14	126	2.5	6474	4 US-09-949-016-1717	Sequence 1717, Ap
15	122.5	2.4	1463	4 US-09-270-767-15085	Sequence 15085, A
c 16	122.5	2.4	31826	4 US-09-902-540-1256	Sequence 1256, Ap
17	121.5	2.4	13977	3 US-09-484-9708-60	Sequence 60, Appl
c 18	121.5	2.4	24333	4 US-09-639-207-9	Sequence 9, Appl
19	119.5	2.4	3545	3 US-08-885-291-54	Sequence 54, Appl
20	119.5	2.4	3545	3 US-09-496-672-54	Sequence 54, Appl
21	119.5	2.4	5715	3 US-09-107-847-1	Sequence 1, Appl
22	119	2.4	580073	4 US-08-545-528D-1	Sequence 1, Appl
23	118.5	2.4	5391	4 US-08-811-519-2	Sequence 2, Appl
24	118.5	2.4	5693	4 US-09-262-537-19	Sequence 19, Appl
25	118	2.3	1380	3 US-09-319-303-1	Sequence 1, Appl
26	118	2.3	2033	4 US-09-270-767-14130	Sequence 14130, A
27	118	2.3	2418	4 US-09-614-221A-495	Sequence 495, App
c 28	118	2.3	16047	4 US-09-902-540-1136	Sequence 1136, Ap
29	118	2.3	36820	4 US-09-949-016-16665	Sequence 16665, A
30	117.5	2.3	1794	4 US-09-949-016-2105	Sequence 2105, Ap
31	117.5	2.3	4926	2 US-08-853-310-1	Sequence 1, Appl
32	117	2.3	1158	4 US-09-248-796A-3462	Sequence 3462, Ap
33	117	2.3	2729	1 US-08-412-431-2	Sequence 2, Appl
34	117	2.3	2729	1 US-08-623-679-2	Sequence 2, Appl
35	117	2.3	2729	3 US-08-933-774-2	Sequence 2, Appl
36	117	2.3	2729	3 US-09-181-030-2	Sequence 2, Appl
37	117	2.3	2729	3 US-09-534-242-2	Sequence 2, Appl
38	117	2.3	2729	3 US-09-454-854-2	Sequence 2, Appl
39	117	2.3	2729	3 US-09-164-671-2	Sequence 2, Appl
40	117	2.3	2729	4 US-09-182-113-2	Sequence 2, Appl
41	117	2.3	2729	4 US-08-862-442-2	Sequence 2, Appl
c 42	116.5	2.3	1859	4 US-09-489-039A-3406	Sequence 3406, Ap
43	115.5	2.3	6671	1 US-08-280-443-1	Sequence 1, Appl
44	115.5	2.3	6671	1 US-08-457-459-1	Sequence 1, Appl
45	115.5	2.3	6671	1 US-08-555-678-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-290
; Sequence 290, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_Fl_genes Version 1.0
; SEQ ID NO 290
; LENGTH: 2523
; TYPE: DNA

Later

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (81)..(2138)
US-09-620-312D-290

Alignment Scores:

Pred. No.: 0 Length: 2523
Score: 3676.00 Matches: 692
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 3
Query Match: 73.02% Indels: 2
DB: 4 Gaps: 2

US-09-671-687a-3 (1-949) x US-09-620-312D-290 (1-2523)

QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrp 274
DB 45 GGAATAAAGCTTTTATATTTTGTGTGGGACATGGATAACCTATTGGCAACCTGG 104
QY 275 AspGlyArgPheAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeu 293
DB 105 GATGAAGATTGATGGAGTGACCTTTGTAGTTTGGGTGTGGTGAAGTACAAATCTA 164
QY 294 LeuHisIleAsnAspIleIleProGluSerValThrGlnGluArgProProLysLeu 313
DB 165 TTGCACATCAATGATATCATCCAGAGAGTGTGACGAGAAAGGAGGCTCCCAAACTT 224
QY 314 AlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAla 333
DB 225 GCCTTTATGTCAAGAGGTGTGGGACCAAGGTTTCATCCAGTCATAATAAACAAGGCT 284
QY 334 ThrGlySerThrSerAspProGlyValAsnArg---ArgSerGluLeuPheTyrThrLeuAsn 352
DB 285 ACAGGATCTACCTCAGACCTCGAATAGAAACACAGATCTGAATATTTTATACCTTAAAT 344
QY 353 GlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu 372
DB 345 GGGTCTCTGTGACTCACACACACCAATCCAAATCAAAATATACATGGTACATTGATGA 404
QY 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSer 392
DB 405 GTTGACAGAGACCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCA 464
QY 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
DB 525 CCATTTCAGTCTCACCAAGATGCCATACCAATGGGAAGTATTGGCCACAGTCCACTTTCT 584
QY 433 LeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProPro 452
DB 585 CTGTACGCCAGTCTGTAATGGAAGAGCTAAACACATGCCGCCCTGCCAAGAGAGTCCACCC 644
QY 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
DB 645 TTGGCCATGCCCTCGGGAACCTACATGGTGTAGAAAGTGGGCTCATTTGGCTGAAGTTAA 704
QY 473 GluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGlu 492
DB 705 GAGAACCTCTCTTATGGGTAATCCGTGTGATCGGTGAGCCAGCACGAGACTGAATGAA 764
QY 493 ValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArg 512
DB 765 GTGCTCGCTGGACTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGA 824
QY 513 GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCys 532
DB 825 GGCACCTCGGTATTTACCTGTGCCCTGGAAGAGGCGCTGTTGTGAACACTGAAGAGCTGC 884
QY 533 ArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552

DB 885 AGGCCTGACTCTAGGTTTGCATCATTCAGCCGGTTTCCAATCAGATTGACCCCTGTAAC 944
QY 553 SerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrProProLysMet 572
DB 945 TCTTTAGCAATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAANAATG 1004
QY 573 GluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsn 592
DB 1005 GAAAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGGATCCAGGGTCAATCAAT 1064
QY 593 SerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThr 612
DB 1065 TCTTGTACTTACTTACTCAACCTTATTTCTGCTTATTTTCTTTTCTGCTGTCGACACT 1124
QY 613 ValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 632
DB 1125 GTGTACTTACTAGACCCAAAGAAAGACGATGTAGAATATATATAGTGAACCCAGAGCTA 1184
QY 633 LeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIle 652
DB 1185 CTGAGGACAGAAATTTCTTAATCTCTGAGAATATATGGATATGTGTGTGCACAAAAATT 1244
QY 653 MetLysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGlu 672
DB 1245 ATCAAACTCAGGAAAAATCTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAA 1304
QY 673 LysAspProGluLupheLeuAsnIleLeuPheHisIleLeuArgValGluProLeu 692
DB 1305 AAAGATCCTGAGGAATTTCTTGAATATTTCTGTTTCATCATATTTTAAAGGTAGAACCTTTG 1364
QY 693 LeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMet 712
DB 1365 CTAAATAATAGATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATG 1424
QY 713 GluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIle 732
DB 1425 GAAAAAATGAGAAAGTTGGCGTTCGCCAAATTCAGCAGTTGTAGAAATGGTCTTTTATC 1484
QY 733 AsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPhe 752
DB 1485 AACAGTAACCTGAAATTTGCAGAGGACCATCATGCTGATTTTCAGATGCCTCGATTT 1544
QY 753 GlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAsp 772
DB 1545 GGAAGAGACTTTAAACTATTAAATAAATTTTCTCTCTGGAATTAATATAACAGAT 1604
QY 773 LeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCys 792
DB 1605 TTACTTTGAAGACACTCCAGACAGTCCGGATATGTGGAGGCTTGCATATGTATGAGTGT 1664
QY 793 ArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThr 812
DB 1665 AGAATAATGCTACACCATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCC 1724
QY 813 CysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSer 832
DB 1725 TGCAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATAACCCAGTGTCA 1784
QY 833 LeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMet 852
DB 1785 CTTCCCAAGACTTACCCGACTGGGACTGGAGACACGCTGCATCCCTTGCACGAATATG 1844
QY 853 GluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGly 872
DB 1845 GAGTTATTTGCTGTCTCTGCTAGAAACAAAGCCACTATGTTGCTTTTGTGAAGTATGGG 1904
QY 873 LysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsn 892
DB 1905 AAGACAGATTCGCTGGCTCTCTTTTGACAGCATGGCCGATCGGGATGGTGGTCAAGAT 1964
QY 893 GlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSer 912

Db 1965 GGCTTCAACATTCCTCAAGTCACCCCATGCCAGAAAGTAGGAGGTACTTGAAGATGTCT 2024
Qy 913 LeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeu 932
Db 2025 CTGGAAGACCTGTCATTCCTTGACTCCAGAGAAATCCAGAGGCTGTGCACGAAGACTGCTT 2084
Qy 933 CysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2085 TGTGATGATATATGTGCAATGTACACAGATCCCAACAAATGAGTTGTACAAA 2135

RESULT 2
US-09-646-403-3
; Sequence 3, Application US/09646403
; Patent No.: 6734174
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: KOVALENKO, Andrei
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR
; TITLE OF INVENTION: AND OTHER PROTEINS
; FILE REFERENCE: WALLACH-27
; CURRENT APPLICATION NUMBER: US/09/646,403
; CURRENT FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: IL 123758
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: PCT/IL99/00158
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: IL 126024
; PRIOR FILING DATE: 1998-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2116
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (691)..(691)
; OTHER INFORMATION: n is unknown.
US-09-646-403-3

Alignment Scores:
Pred. No.: 1,28e-266 Length: 2116
Score: 2468.00 Matches: 524
Percent Similarity: 91.21% Conservative: 5
Best Local Similarity: 90.34% Mismatches: 33
Query Match: 49.03% Indels: 22
DB: 4 Gaps: 6

US-09-671-687A-3 (1-949) x US-09-646-403-3 (1-2116)

Qy 387 AspPheAspArgSerProProLeuGlnProProProValAsnSerLeuThrThrGlu 406
Db 15 GACTTTGACCGTCTTCCACACCACTCCAGCCCTCCTCTGTGACTCACTGACCACCGAG 74
Qy 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426
Db 75 AACAGATTCCTCTTACCATTCAGTCTCACCAGATGCCCAATACCAATGGAAGTATT 134
Qy 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuAsnThrAlaPro 446
Db 135 GGCACAGTCCACTTCTCTGTGAGCCCTCTGTAATGGAAGAGTAAACACTGACCCC 194
Qy 447 ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly 466
Db 195 GTCCAAGAGAGTCCACCTTGGCCATGCTCTCTGGGAATCTACATGGTCTAGAAAGTGGGC 254
Qy 467 SerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGln 486
Db 255 TCATTGGCTGAAGTTAAGAGAACCTCTCTCTATGGGGTAATCCGTTGGATCGGTGAG 314
Qy 487 ProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCys 506
Db 315 CCACCAGGACTGAATGAAGTGTGCTGCTGAGTGGAACTGGAAGTGAAGTGTGCTGAGGCTGT 374

Qy 507 ThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPhe 526
Db 375 ACGGATGGAAACCTTTCAGAGCACTCGGTATTTCCCTGTGCCCTGAAAGAGCGCTGTTT 434
Qy 527 ValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsn 546
Db 435 GTGAAACTGAGAGAGCTGCAGGCTGACTCTAGGTTTGCATCATTTGCAGCCGGTTCCAA 494
Qy 547 Gln-11eGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluG 566
Db 495 CAAAGATTGAGCGCTGTAACTCTTTAGCATTTGGAGGCTACTTAAAGTGAAGTAGT-GAAGA 553
Qy 566 uAsnThrProLysMetGluLysGluGlyLeuGluIleMetIleGly----LysLysLys 585
Db 554 AAATACTCCACCAAAATGGAAAAGAAAGCTTGGAGATAATGATGGGGAAGAAAGAAA 613
Qy 585 sGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPhe-CysLeuPhe- 604
Db 614 GGCATCCAAGGGTCAATTACAAATCTTGKTACTTAGACTCAACCTTATCTKGCTTATTK 673
Qy 605 AlaPheSerSerValLeu-AspThr---ValLeuLeuArgProLysGluLysAsnAsp-- 622
Db 674 GCTTTTAGTTCGTCTCTNGGACACTGCTGTTACTTTTAGACCCCAAGAAAAGAAACGAT 733
Qy 623 -ValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValLeuProLeuAr 642
Db 734 GTTAGAATATTTWTWKWMMACCCCAAGAGCTACTGAGGACAGAAATGTTAATCTCTGAG 793
Qy 642 gIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlyLeuVa 662
Db 794 AATATATGGAATATGTGTGCCCAAAAATTTATGAAACTGAGGAAAATACTTGAAGAGGT 853
Qy 662 lGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIleLe 682
Db 854 GGAGGCTGCATCAGGATTTACCTCTGAGAGAAAAGATCTCTGAGGAATTTCTGAATATCT 913
Qy 682 uPheHisHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysVa 702
Db 914 GTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGT 973
Qy 702 lGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProTh 722
Db 974 ACAAGATTGTTACTTCTATCAAAATTTTATGAAAAAAATGAGAAAGTTGGCGTTCAC 1033
Qy 722 rIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaPr 742
Db 1034 AATTGACAGTGTGTAGATGGTCTTTTATCAACAGTAACTGAAATTTGAGAGGCACC 1093
Qy 742 oSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysIle 762
Db 1094 ATCATGTCTGATTATTTCAGATGCTCGATTGGAAAAGACTTTAAACTATTTAAAAAATT 1153
Qy 762 ePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysAr 782
Db 1154 -TTTCCTTCTCTGGAATTAGATATAACAGATTTTACTTGAAGACAC-CCCAGACAGTGC 1211
Qy 782 gIleCysGlyLysLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSe 802
Db 1212 GATATGAGGAGGCTTGCAATGATGAGTGTA- GAATGCTACGACCATCCGACACCCAG 1270
Qy 802 rAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLy 822
Db 1271 C---TGAAAAACAAGCAGTTTGTAAAACTGCAACACTCAAGTCCACCTTCATCCGA 1327
Qy 822 sArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAsp---TrpAs 841
Db 1328 GAGGCTGAATCATATAATAAACCCAGTGTCACTTCCCAAGACTTACCCGACTGGAGA 1387
Qy 841 pTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleG 861
Db 1388 TTGGAGACCGGCTGCATCCCTTGGCCAGAAATATGGAGTTATTTGCTTCTCTCATAGA 1447

Qy 861 uThSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaThrPhePhe 881
Db 1448 AACAGCCACTATGTTGCTTTTGTGAAGTATGGAGGACGATTCCTGCTGCTCTTCTT 1507
Qy 881 e-AspSerMetAlaAsp-ArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
Db 1508 TGGACAGCATGCCATCCGGATGGTGTGAGATGGCTCAACATTCCTCCCAAGTCMCC 1567
Qy 901 ProCysProGluValGlyGlnTyrLeu-LysMetSer-LeuGluAsp---LeuHisSerL 919
Db 1568 CMTGCCAGAGTAGGAGTACTTGGAGATGCTCTCTGGAAGACCTGSAWTCCTT 1627
Qy 919 euAspSerArgArgile-GlnGlyCysAlaArgArgLeuLeuCysAspAla-TyrMetCy 938
Db 1628 GGACTCCAGGAGAAATCCAGGCTGTGCACGAGACTGCTTTGTGATGCCATATATGTG 1687
Qy 938 s-MetTyr-GlnSerProThrMetSerLeuTyrLys 949
Db 1688 CCATGTACCCAGAGTCCCAACATGAGTTGTACAAA 1723

RESULT 3

US-09-513-999C-20085
; Sequence 20085, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 20085
; LENGTH: 110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-20085

Alignment Scores:
Pred. No.: 4.2e-12 Length: 110
Score: 185.00 Matches: 36
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.68% Indels: 0
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-513-999C-20085 (1-110)

Qy 595 TyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeu 614
Db 1 TACTTAGACTCAACCTTATTCGCTTATTTGCTTTTGTGTTCTGTCGACACTGTGTTA 60
Qy 615 LeuArgProLysGlnLysAsnAspValGluTyrTyrSerGluThrGln 630
Db 61 CTTAGACCCCAAGAAAGACGATGAGTAATATTATAGTGAACCCAA 108

RESULT 4

US-09-220-132-79
; Sequence 79, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303

; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 5857
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-79

Alignment Scores:

Pred. No.: 4.31e-05 Length: 5857
Score: 151.00 Matches: 107
Percent Similarity: 31.58% Conservative: 61
Best Local Similarity: 20.11% Mismatches: 184
Query Match: 3.00% Indels: 180
DB: 4 Gaps: 23

US-09-671-687A-3 (1-949) x US-09-220-132-79 (1-5857)

Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 277 GAGACTCAGGAGGAATTTGTGATGACTTTCGAGTGGGAGGAGCGAGTTTGGTGAATGGA 336
Qy 239 -----GlyGluThrIleGluSerGlyThrValIlePhe 249
Db 337 AATAAGCCTGGATTATCCAGTTTCTTGAGAAACC----- 372
Qy 250 CysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsn 269
Db 373 ---CAGTTTGCACAGCCAG-----TGGCTGGAATTTTGTAGATGAA 414
Qy 270 ProIleGlyAsnTrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu 289
Db 415 CCATAGGCAAGAACGATGCTCGGTGGCAGGAGTTCGG---TATTTCCAGTGTGAACCT 471
Qy 290 SerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArg 309
Db 472 -----TTAAGGGCATATTT-----ACCGA 492
Qy 310 ProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHisAsn 329
Db 493 CCTTCAAGTTA-----ACAAGGAGGTG-----CAAGCA 522
Qy 330 LysProLysAlaThrGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyr 349
Db 523 GAAGATGAAGCTAATGGCTTGCAGACAAACCCGCTCCCGAGCTACTTCACCGCTGTGC 582
Qy 350 ThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyr 369
Db 583 ACTTCTACGGCCAGCATGGTGTCTTCTCCCTCCACCCCTTCAAC----- 630
Qy 370 IleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAsp 389
Db 631 ATCCCTCAGAAACCATCACAGCCAGCA----- 657
Qy 390 ArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPhe 409
Db 658 ---GCAAGGAACCTTCAGCTACGCTCCGATCAGCAACCTTACA----- 699
Qy 410 HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSer 429
Db 699 ----- 699
Qy 430 ProLeuSerLeuSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGlu 449
Db 700 -----AAAACTGCCAGTGAATCTATCTCAACCTTTTCAGAGGTGGCTCAATCAAGAAA 753
Qy 450 SerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAla 469
Db 754 -----GGAGAAAGAGAGCTCAAAATCGGAGACAGAGTA 786


```
QY 470 GluValLysGluAsnProPheTyrGlyValIleArgTrpIleGlyGlnProProGly 489
Db 787 TTGGTTGGGCGACTAAGGCT-----GGTGTAGTCCGGTTCTTGGGGAGACCGACTTT 840
QY 490 LeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGly 509
Db 841 GCCAAGGGGGAGTGGTGGCGTGGAGTTA--GATGAGCCACTTGGGAAGAATGATGCG 897
QY 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
Db 898 GCTGTGTGGCAACAAGGATTTTTCAGTGTCAACCCAAATATGCTGTGCTGCT----- 951
QY 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
Db 952 -----CTGTCCAC 960
QY 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrPro 569
Db 961 AAAGTTTACCAAGATTGGCTTC-----CCTTCCACTACACCA 996
QY 570 ProLysMetGluLysGluGlyLeu---GluIleMetIleGlyLysLysLysGlyIleGln 588
Db 997 GCCAAGCCCAAGCCCAACGAGTGGCGAGTGTATGCGCAGCACGTCGCCGCGAGCTGAAG 1056
QY 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608
Db 1057 CCGAGCCCTTCGCTTCCTCCCTCAGCTCC-----ATGAGCTCA 1095
QY 609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628
Db 1096 GTGGCTCTCTGTGAGCAGCGGCC-----AGTCGG 1128
QY 629 ThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCys 648
Db 1129 ACAGGACTATTGACTGAAACCTCTCCCTTACGCCAGGAAGATCTCCGCTACCACTGCC 1188
QY 649 AlaThrLysIleMetLys-----LeuArgLysIle 658
Db 1189 CTCAGGAGGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1248
QY 659 LeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPhe 678
Db 1249 CTGGAGAGGCGGAGGTGGCCCAAGGCC---ACGAGCCAGTGGGGAGATAGAGCAGGAG 1305
QY 679 LeuAsnIleLeu-----PheHisIleLeu----- 687
Db 1306 CTAGCTCTGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1365
QY 688 -----ArgValGluProLeuLeuLysIle 695
Db 1366 CAGCTCGCAACAATGTTGGAAGCTGCTGACAGGAGAGGTTGAGCTTCTCAACCAAGCTT 1425
QY 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPhe 707
Db 1426 GAAGAGGAGAAAGGAAGGTTGAGGACCTTCAGTTC 1461
```

RESULT 5

```
US-09-949-016-1895
; Sequence 1895, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```



```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1895
; LENGTH: 2825
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1895

Alignment Scores:
Pred. No.: 0.000251 Length: 2825
Score: 139.00 Matches: 131
Percent Similarity: 33.06% Conservative: 74
Best Local Similarity: 21.13% Mismatches: 232
Query Match: 2.76% Indels: 184
DB: 4 Gaps: 28
```

US-09-671-687A-3 (1-949) x US-09-949-016-1895 (1-2825)

```
QY 51 GlnAspArgSerValGlyHisSerArgIleProSerAlaLysGlyLysAsnGlnIle 70
Db 17 CAAGCCAGAGAGCGGAGGAAATGAAATTCCTCAATGCC----- 55
QY 71 GlyLeuLysIleLeuGluGlnProHisAlaValLeuPheValAspGluAspValValGlu 90
Db 56 -----CTCTCTAGCTGGTTGACCTCCGCTGTGATGCAAC----- 94
QY 91 IleAsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGluGluArgPheSer 110
Db 95 -----TCAAGCCGCGCTCTCA 109
QY 111 LeuPhe-----LysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspVal 126
Db 110 CTGTGGCTTGTGCTCCGGGAAATGGAACCCAGCTTGGGACAGGATGGATGATTTG 169
QY 127 GlyCysProValLysValGlnLeuArgSerGlyGluLysPheProGlyValValArg 146
Db 170 ACTTGCTTTTGGCGTCTCCCGAGCCCTGTGGAGCCAG----- 208
QY 147 PheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeu 166
Db 209 -----GCCTCTTGGAGCATCTTTGGGGCTGACGCA 238
QY 167 LeuGluGluArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPhe 186
Db 239 GCGGAGGTTCGCGGCACACGTCGCCACTCCCGCAG----- 274
QY 187 GlnCysAspLysAspCysGlyPheValAlaLeuAspLysLeuGluLeuGluAspAsp 206
Db 275 -----GAGGCTGCC 283
QY 207 AspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuPro 226
Db 284 ATGCCCCACATTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 343
QY 227 Pro-----LeuGluIleAsnSerArg-Valse 235
Db 344 GCGGCGCAAGTAAGTCTCTACCGCAGGAGTATCTTGAGTCCCACTCCACGATGT 403
QY 235 rLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProG 255
Db 404 CTGACTGGGGATGCCACTTTCACCAAGAGGAGAAAC-----GACAA-AAACCTGG 453
QY 255 YLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAs 275
Db 454 CMACAGAGT-----TCACAGTCCCAAGAGGCTTTCTC 489
QY 275 pGlyArgPheAspGlyValLeuCys-SerPheAlaCysValGluSerThr-----IleL 293
Db 490 ACCGACACTTGAAGGTGTCACCTGCTCCCTGACATCTGTGGACCCCGGGGGCACATCA 549
QY 293 euLeuHisIleAsnAspIleIleProGlu---SerValThrGlnGluArgProProL 312
Db 550 TTGACCTGGTGAATGATCCAGCTGCCAGACATCATCATCTCAGAGGAGGAGGAGGAGGAGGAG 609
```

QY 312 ysLeuAlaPheMet-----SerArgGlyValGlyAspLysGly 325
 Db 610 ACCTGGCGCTGCTGGAGAGCAAGTTGGTGAGTGGCGATTCCTGACCGCGCTGGGA 669
 QY 325 erSerSerHisAenLysProLysAlaThrGlySerHisAenArgArg 345
 Db 670 GGAAGTCCAGGACGCCCGGAGACTCCCCATCAGCTGTTCCCGCAACCTCAGCCCCA 729
 QY 345 erGlu-LeuPheTyrThrLeuAenGlySerSerValaspSerGlnProGlnSerLysSer 364
 Db 730 GCGCTTCTCTACATCCTCTCGAAGCACTACTTACAGTCCCAACCCCGCCAGAGGG 789
 QY 365 LysAenThrTrpTyrLeuAspGluValAlaGluAspProLa-----Lys 379
 Db 790 GATCAGGCC-----GACGTCTCTTCACTCAGCTCAGCTGGCGAGCTAACGTCGCCAAA 840
 QY 380 SerLeuThrGluLeuSerThrAspPheAspArgSer-----SerPro 393
 Db 841 GGCTAGCTACGAGGAGCAATGACAGAGAAAGTGTCTCAGGCGAGCTGGCTCTCT 900
 QY 394 ProLeuGlnProProValAen-----SerLeuThrThrGluAenArg 408
 Db 901 -----CGTCTCTCCAGTTGAGAAGTCAAAGAGATTGCATAGAACAAAGGAAAC 954
 QY 409 PheHisSerLeu-----ProPheSerLeuThrLysMetProAenThr 422
 Db 955 TTGATCCCTCCAGTACCCGAGACACACCAAGGCGCTAGCTCTGTTTACAAACAGC 1014
 QY 423 AsnGlySerLeuGly---HisSerProLeuSerLeuSerAlaGlnSerValMet----- 439
 Db 1015 AGTGGGAAATGGCCCTGAAACAGCCCTCAGCTGGCCCGCTGGAGCGAGCTGGGGAAG 1074
 QY 440 -----GluGluLeuAsnThrAlaProValGlnGluSer 450
 Db 1075 CAGCTTTGAAAACGGGCTGGAGGGCAGCCCTTCGCCGAGAGTCCAAACCCAGGATGCG 1134
 QY 451 -----ProProLeuAlaMet-----ProGlyAenSerHisGlyLeu 463
 Db 1135 GCAGGAGTGGTCCCGCAGCTCCCGGGGAGAGCCAGCTGGAGAGCGATGGGGCCC 1194
 QY 464 GluValGlySerLeuAlaGluValLys-----GluAsnProProPheTyrGlyVal 480
 Db 1195 GAGCTGGCTCCAAAGCTGAGCTTCCACCACCTGTGTCCCGGCGCCCGCTGCGAGGG 1254
 QY 481 IleArgTrpIleGlyGlnPro-----ProGlyLeuAsnGluValLeuAlaGly 496
 Db 1255 CTCTCTGGGACAGTGGCCCTGAAAGACCTGGCCCCCGCTGCAAAAGTGTTCGCCAAG 1314
 QY 497 LeuGluLeuGluAspGlu-----CysAlaGlyCysThrAspGlyThrPheArg--- 512
 Db 1315 CTGCCACTGGCAGAGGAGAAACGTTTTCAGGCAAGCGCGCGCAAGCTGGCCCAAG 1374
 QY 513 -----GlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLys 530
 Db 1375 GCCCTGGTCTCAAAGACTTTCAGATACAAGTGCAGCCGCTGCGGATGCAGAACTGACC 1434
 QY 531 SerCysArgProAspSerArgPhe-----AlaSerLeuGln 542
 Db 1435 AAGCTCCGAGGAGGACATCCTGATGAGAAATCAGAACTTAGTGGGGCTCAAGCTTCCA 1494
 QY 543 ProValSerAenGlnIleGluArgCysAenSerLeuAlaPheGlyGlyTyrLeuSerGlu 562
 Db 1495 GACCTTAGTGAGCAGCTGACGAGGAAAAGGGCTTCT-----TCTGAACCTCTCCCA 1548
 QY 563 ValValGluGluAenThrProProLysMetGluLysGluGlyLeuGluIleMet 580
 Db 1549 GCTATTGAGGAA-----GAAGAGTCAAAGAGTGGCTTAGATGTGTCATG 1590

RESULT 6
 US-09-949-016-5779
 ; Sequence 5779, Application US/09949016
 ; Patent No. 6812339

; GENERAL INFORMATION:

```

; ORGANISM: HUMAN.
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5779
; LENGTH: 3130
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5779

```

Alignment Scores:

Pred. No.:	0.000304	Length:	3130
Fragment Score:	139.00	Matches:	102
Score:	33.40%	Conservative:	56
Percent Similarity:	21.56%	Mismatches:	176
Best Local Similarity:	2.76%	Indels:	140
Query Match:	4	Gaps:	20
DB:			

US-09-671-687A-3 (1-949) x US-09-949-016-5779 (1-3130)

Qy	309	ArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHis	328
Db	22	AAGCCCC-----GGCCGTGGGGAAG-----CAC	48
Qy	329	AsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArgArgSerGluLeuPhe	348
Db	49	TCACGCCCCATGGCCGGACATCTACTGGTCAGCTTCATCTCGGCGCGGTGCCGCT	108
Qy	349	TyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrp	368
Db	109	AGCTCCAAGGAAGGCTCCCACTGCACAAACAGTCATCTGACCCCTCTCTCCCG---	165
Qy	369	TyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPhe	388
Db	166	CGCGCAGCTGCTGCCCCCGAAGCCGGGCCCAAGGCGCGAAGTGGGGATGACATC	225
Qy	389	-----AspArgSerSerProProLeuGlnProProValAsn	401
Db	226	CTGGGGACTTTGTGGTGGCGAGCGGGTGTGGGTGAACGGCGTGACCGCGTGGTG	285
Qy	402	SerLeuThrThrGluAsnArgPhe-----	409
Db	286	CAGTATCTGGAGAGACGCAATTCCACCGGCCAGTGGGCTGGCGTGGTGGACGAC	345
Qy	409	-----	409
Db	346	CCGTGGGCAAGATGATGGCGGTGGGCGCGCTGCGCTACTTCAGTGC CGCGCCCTC	405
Qy	410	-----HisSerLeuProPheSerLeuThrLysMetProAsnThr-----AsnGly	424
Db	406	CAGGGTATCTTCACGGCGCCCTCCAGCTGACCCCGCGACGCCACGCGCGAGGCTCGGGG	465
Qy	425	SerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuAsnThr	444
Db	466	AGTGATGCCCACTCCGTGGAGTCGTGACTGCCCAAGACCTGTCTATTGCAATTCGGGCACG	525
Qy	445	AlaProValGlnLeuSerProLeu-----	453
Db	526	GCC-----ACGCCCCCGCTGACCAAGCCGCTATCCCCCTCGGGAGAGCGTC	573
Qy	454	-----AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGlu	470

574 CTCAACAGCTCCGTGAAGACCGGACAGAGTCCGGATCCAACTCTCAGACAGCGGCTCT 633
QY
471 ValLysGluAsnProPheTyr----- 478
Db
634 GTGAAGCGGGCGAAAGAGACCTGCGCCTGGGGACCGGCTGCTGGTGGCGGAGCAAG 693
QY
479 ---GlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeu 497
Db
694 ACTGGCGTGGTGGTGGTGGGAGAGACAGACTTTGCCAAGGGCGAGTGGTGGCGG 753
QY
498 GluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPhe 517
Db
754 GAGCTG---GACGAGCCCTTGGAGAGATGATGGGGCGTGGCGGACACAGTACTTC 810
QY
518 ThrCysAlaLeuLysLysAlaLeuPheValLysLeu----- 529
Db
811 CAGTGGCCACCAAGTTTGGTCTCTTCGCGCCCATCCAAAGTGATCGGTATCGGCTTC 870
QY
530 -----LysSerCysArgProAspSerArgPheAla----- 539
Db
871 CCATCTACAGCCCGCAGCCAGGCAAGCAAGAGACCAAGCGTATGGCCCATGGTGTGCAGCA 930
QY
540 ---SerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAla-----Phe 556
Db
931 CTGACCCACAGTCCACAGATTCCTCCATCAGCTCCGTCAGCTCTGTGGCTCTCCGTC 990
QY
557 GlyGlyTyrLeuSerGlu-----ValValGluGluAsnThrProProLysMetGlu 573
Db
991 GGGGTGGCGCCAGCGCAGTGGCTGCTCAGGAGACCTCTTCACGCTAGCGCCGCAAG 1050
QY
574 LysGluGlyLeuLysMetIleGlyLysLysLysGlyIleGlnGlyHis-TyrAsnSe 593
Db
1051 APTCTGGGACCAACCGCCTTGAGGAGGACCTGAAGGAGAGCAGCAGCAGCATTTGAGCAG 1110
QY
593 rCysTyrLeuAspSerThr-----LeuPheCysLeuPheAlaPheSerSerValLe 610
Db
1111 CTGCTGGCTGAACGAGACCTGGACCGGCTGAG-GTGGCCAGGCCACCAAGCCACATCTG 1169
QY
610 uAspThr-----ValLeuLeuArgProLysGluLysAsnAspValGluTyrTy 626
Db
1170 CGAGGTGGAGAGGAGATTGCCCTGCTCAAGGACACAGCATGAG-----CAGTATGT 1220
QY
626 rSerGluThrGlnGlu---LeuLeuArgThrGluIleValAlaAsnProLeuArgIleTyrGl 645
Db
1221 TGCAGAAACCGAGGAGAGCTGAGCGAGCGCGGCTGCTC----- 1260
QY
645 yTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuLysValGluAla 665
Db
1261 -----GTGAGAGCGTGGGAAA-----GAGAAGGTGACCTGTC 1295
QY
665 sSerGlyPheThrSerGluLysAspProGluGlu 677
Db
1296 CAACAGCTGGAGGAGGAGGAGGAGGAGTGGAGGAT 1332

RESULT 7

US-09-399-913-56
; Sequence 56, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Huai-Ping
; APPLICANT: An, Wendian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USN 60/109,333
; EARLIER FILING DATE: 1998-11-20

EARLIER APPLICATION NUMBER: USN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 2643
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(801)
US-09-399-913-56
Alignment Scores:
Pred. No.: 0.00627 Length: 2643
Score: 135.00 Matches: 131
Percent Similarity: 31.65% Conservative: 51
Best Local Similarity: 22.78% Mismatches: 172
Query Match: 2.68% Indels: 222
DB: 3 Gaps: 30

US-09-671-687A-3 (1-949) x US-09-399-913-56 (1-2643)

QY 169 GluGlyArg-----GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGln 184
Db 201 CAAGGACAGGACCGCTGGTGGCCAAAGAAATTCGACGCGCTGTAGAGAGGC----- 254
QY 185 LeuPheGlnCysAspGlu-AspCysGlyPheValAlaLeuAspLysLeuGluLeuIleGl 204
Db 255 -----TGTCCTCCTAAAGTCACACTACCCAA 296
QY 204 uAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGl 224
Db 297 CTATGCAACAGTC-----CCAGGCAATCTCATGCTCAGCGC 332
QY 224 uLeuProProLeuGluIleAsnSerArgValSerLeuLysGlyGlyGluThrIleGluSe 244
Db 333 GCTG---GGCTCGCTCTAGGACACGAGTGTCTCTCGATGGCCAGAGAGC----- 381
QY 244 rGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGl 264
Db 382 -GGACGCTGAGTTCCTGGGGACACCC-----GAGTTCGACAGTGGCCAGTGGTGG 434
QY 264 yValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal----- 281
Db 435 CGTGGAGCTAGATGAACCGGAGGCAAGACGACGCGGAGCGTGGGGGTGTCCGGTACTT 494
QY 282 -LeuCysSer-----PheAlaCysValGluSerThrIleLeuLeuHisI 296
Db 495 CATCTGCCCTCCCAAGCAGGGTCTTTTGTATCTGTGTCCAAG-----GTCCTCAAGGC 548
QY 296 eAsnAspIlelleProGluSerValThrGlnGluArgProProLysLeuAlaPheMe 316
Db 549 AGTGGATGACCCCTCATCTGTACTCCACGCCCGCAGCTCCCGGATGGACTTC-- 606
QY 316 tSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySe 336
Db 607 -TCCCGTGTAAACGGGC---AAAGCGGAGGAAACACAAA----- 642
QY 336 rThrSerAspProGlyAsnArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
Db 643 -----GGGAAGAAAGTCCCA----- 660
QY 356 lAspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAs 376
Db 660 ----- 660
QY 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGl 396
|||||

```
Db 661 -----TCTTCCCCA----- 669
Qy 396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerIe 416
Db 669 -----
Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436
Db 670 -----TCTCTGGC-----AGCGTGA 686
Qy 436 n-SerValMetGluGluLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetP 456
Db 687 GCAGCGTG-----AAGGGCCAAAGCTGAAG----- 712
Qy 456 roProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProp 476
Db 713 -----TTGGAGACCAAGTCTTGTGGCAGGCCAACA----- 745
Qy 476 roPheTyGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaG 496
Db 746 -----GGGATTGTGGCTTCTATGGGAAGACAGACTTGTCTCCAGGTACTGGTATG 797
Qy 496 lyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgT 516
Db 798 GCATTGAACGTG---GACCAGCCACACGGCAAGCATGACGGCTCTGTGTTCGGTGTCCGGT 854
Qy 516 yrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspS 536
Db 855 ACTTTACCTGTGCCCGAGGCGAGGGGTCTTTGCA-----CAGCAT 896
Qy 536 erArgPhe-----AlaSerLeuGlnProValSerAsnGlnIleGluArgCysA 552
Db 897 CTCGTATCCAGAGGATTGTGTGGATCCACTGATCCCTCGGACAGACTGTGGAGCAAA-A 955
Qy 552 snSerLeuAlaPheGlyGlyTyLeuSerGluValValGluGluAsnThrPro----ProL 571
Db 956 AAGGTGCATCAAGTGACATGACACACAGCCCAACAGCACCTTCACACACAGTCGGACCCA 1015
Qy 571 yMetGluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHis- 590
Db 1016 AA-----GGACATT 1024
Qy 591 --TyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPhe----- 606
Db 1025 GCATCAGAACTCATCTCCAGTTACTTCTCTGTCTGTGGTGTCTTCTTGGATGCTGAG 1084
Qy 607 -----SerSerValLeuAspThrValLeu----- 614
Db 1085 GCGGAGATGCGTCTTAGAGACCTGGATACCTGACACAGACAGAGTCCCTCTAGCAT 1144
Qy 615 -----LeuArgPro-----LysG 619
Db 1145 CTCCTGACAAAGGAGACCCAGTCCCTAAGATAGAGATTCCCAAGTACACCTCCAGA 1204
Qy 619 lulyeAsnAspValGluTyTySerGlnGluLeuLeuArgThrGlnIleValA 639
Db 1205 ATAGAAACCCCGTTAGCAGCCCTCATTAAGTCCCATATTAACAGATCTCCAT 1264
Qy 639 snProLeuArgIleTyGlyTyValCysAlaThrLysIleMetLysLeuArgLysIleL 659
Db 1265 GACGACTCCCCAAATACAGACCTCATGTTACCCCAAAAGAGATTCCCTGAGTAGCACCT 1324
Qy 659 euGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheL 679
Db 1325 TCAGGTAGTCCCTGTCCCC-----TACCCCTCAGACAGATTTCCTCCCAATAAACATT 1378
Qy 679 euAsnIleLeuPheHisIleLeuArgValGluProLeu 692
Db 1379 TTCCA-----CATCCCCAAGGATGCTGACCTCTC 1410
```

RESULT 8

US-09-350-614-56

; Sequence 56, Application US/09350614

```
; Patent No. 6689581
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Huai-Ping
; APPLICANT: An, Wendian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP
; CURRENT APPLICATION NUMBER: US/09/350,614
; FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 56
; LENGTH: 2643
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(801)
; US-09-350-614-56
```

Alignment Scores:

Pred. No.:	0.000627	Length:	2643
Score:	135.00	Matches:	131
Percent Similarity:	31.65%	Conservative:	51
Best Local Similarity:	22.78%	Mismatches:	172
Query Match:	2.68%	Indels:	222
DB:	4	Gaps:	30

US-09-671-687A-3 (1-949) x US-09-350-614-56 (1-2643)

```
Qy 169 GluGlyArg-----GlyGlnGlyPheThrAspGlyValTyGlnGlyLysGln 184
Db 201 CAAGGACAGAGCGACCGCTGTGTGCCAAGAAATTGCGGACGCTGTAGAAGAGGC----- 254
Qy 185 LeuPheGlnCysAspGlu-AspCysGlyPheValAlaLeuAspLysLeuGluLeuIleG 204
Db 255 -----TGTGCCACTGTCTGC-----ACCTTCTTAAGTCACACTACCA 296
Qy 204 uAspAspThrAlaLeuGluSerAspTyAlaGlyProGlyAspThrMetGlnValG 224
Db 297 CTATGACACGTC-----CCAGGCAATCTCATGCTCAGCGC 332
Qy 224 uLeuProProLeuGluIleAsnSerArgValSerLeuLysGlyGlyGluThrIleGluSe 244
Db 333 GCTG---GGCCTCGCTTAGGAGACCGAGTCTCTCGATGGCCAGAGACG----- 381
Qy 244 xGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyPheValG 264
Db 382 -GCACGCTCAGGTTCCTGCGGACCCACC-----GAGTTCGCCAGTGGCCAGTGGTGG 434
Qy 264 yValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal----- 281
Db 435 CGTGGAGCTAGATGAACCGGAAGCAAGACGCGGACGCGTGGGGGTGTCGGTACTT 494
Qy 282 -LeuCysSer-----PheAlaCysValGluSerThrIleLeuLeuHisI 296
Db 495 CATCTGCTCCCTCCCAAGCAGCGGTCTCTTTGTGATCTGTGTCCAAAG-----GTCTCAAGGC 548
Qy 296 eAsnAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMe 316
Db 549 AGTGATGACACCCCTCATCTGTACTCTCACGCCCGGACCTCCCGGATGGACTTC-- 606
Qy 316 tSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySe 336
```

Db 607 -TCCCGTGTAACGGGC---AAAGCGCGAGGGAACACAAA----- 642
Qy 336 rThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
Db 643 -----GGGAAGAAGAGTCCCA----- 660
Qy 356 lAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAs 376
Db 660 ----- 660
Qy 376 pProLalysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGl 396
Db 661 -----TCTTCCCA----- 669
Qy 396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416
Db 669 ----- 669
Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436
Db 670 -----TCTCTGGGC-----AGCCTGCA 686
Qy 436 n-SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetP 456
Db 687 GCAGCGTG-----AAGGGGCCAAAGCTGAAG----- 712
Qy 456 roProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProp 476
Db 713 -----TTGGAGACCAAGTCTGTGGCAGCAGAAC----- 745
Qy 476 roPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLeuAlaG 496
Db 746 -----GGGATTGTGGCTTCTATGGGAAGACAGACTTGTCTCAGGTTACTGTGTATG 797
Qy 496 lLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgT 516
Db 798 GCATTGAAGT---GACCAGCCACCGGCAAGCATGCGCTGTGTGCGGTGTCGGT 854
Qy 516 yrPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspS 536
Db 855 ACTTTTACCTGTGCCCGGAGCAGCGGGTCTTTGCA-----CCAGCAT 896
Qy 536 erArgPhe-----AlaSerLeuGlnProValSerAsnGlnIleGluArgCysA 552
Db 897 CTCGTATCCAGAGGATTGTGTGGATCCACTGATCCCTCGAGACAGTGTGGAGCAAA-A 955
Qy 552 snSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluAsnThrPro---ProL 571
Db 956 AAGTGTCATCAAGTGAACATGACACAGCCCAACAGCACCTTCAACAGTCCGAGCCCA 1015
Qy 571 ysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHis- 590
Db 1016 AA-----GGACATT 1024
Qy 591 --TyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPhe----- 606
Db 1025 GCATCAGAGAACTTATCTCCAGGTACTTCTGTCTGTGTGTTCTCTGGATGCTGAGG 1084
Qy 607 -----SerSerValLeuAspThrValLeu----- 614
Db 1085 GCGGAGATCGATCTTAGAGACTGTGATACCTGACACAGACAGAGTCCCTCTAGCAT 1144
Qy 615 -----LeuArgPro-----LysG 619
Db 1145 CTCCTGACACAGGAGACCCCTGACCTACCTTAAGATAGAGATTCCAGTGACACCTCCGA 1204
Qy 619 lLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValA 639
Db 1205 ATAGAACCCTGTTAGCCAGCCCTCGATTACTAGGTCCCATTTATTAACAGATCTCCAT 1264
Qy 639 snProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleL 659
Db 1265 GACGACTCCCAATAACAGACCTCATGTATACCCCAAAAGAGATTCCCTGAGTAGCACCT 1324

Qy 659 euGluLysValGluAlaAlaSerGlyPheThrSerGluLysAspProGluGluPheL 679
Db 1325 TCAGGCTAGTCCCTGTCCCC-----TACCCTCAGACAGATTTCCTCCCAATAAACATT 1378
Qy 679 euAsnIleLeuPheHisHisIleLeuArgValGluProLeu 692
Db 1379 TTCCA-----CATCACCAGGGATGCTGACCTCTC 1410

RESULT 9

US-09-774-528-352
; Sequence 352, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774, 528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 352
; LENGTH: 3245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (402)..(2537)
US-09-774-528-352

Alignment Scores:
Pred. No.: 0.00154 Length: 3245
Score: 133.00 Matches: 168
Percent Similarity: 33.02% Conservative: 110
Best Local Similarity: 19.95% Mismatches: 296
Query Match: 2.64% Indels: 271
DB: 4 Gaps: 39

US-09-671-687A-3 (1-949) x US-09-774-528-352 (1-3245)

Qy 246 ThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyVal 265
Db 119 TCCTAGCCTTCTGTCCAGTCCCTGGCAT-GATCTGATCGCTTTGAAGCAGGTACC 177
Qy 266 AspMetAspAsnProIleGlyAsn-----TTP----- 274
Db 178 CGCTCGGTCTCCCCATCAGCTCCAGGGGCTCTCAGTTGAGTGGTGAGCCGACAGAA 237
Qy 275 -----AspGlyArgPheAspGlyValLeuCysSerPheAlaCys----- 287
Db 238 GCCGGGGGTGAGCTCAAGGGCGGATGGGAGGGGTCTCTGGGGGCGCTTAACCTGCGCACCGC 297
Qy 288 ---ValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGln 306
Db 298 GGCCATCAGCTGGCAGATCAITACAG-AAGTCTCTGTCTCTCTTTCATCTACAGTGC 356
Qy 307 GluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSer 326
Db 357 AGCTGTTTCCAGACGCTTGTCTCCTCAGGTGCG-----GGGAGTGAT 398

```
Qy 327 SerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArgSerGlu 346
Db   : : : : :
Qy 339 CTGATGCCAGGTTTCGAAACTTCTTTCGCTCGGCTCGGTCGCGC----- 449
Db   : : : : :
Qy 347 LeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsn 366
Db   : : : : :
Qy 450 -----TGGATCTCCGGAGGTGGGAGGAGGCTCTCTCTGAGGAGCGGTTCGAGAG 500
Qy 367 ThrTyrTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThr 386
Db   : : : : :
Qy 501 GCGGGGAAATGAGAGGCGCGCGGGGGCTACGAAGCGCTCTTCGAGAGCTGAAGCC 560
Qy 387 AspPheAspArgSerSerProLeuGln-----ProProValAsnSerLeu 403
Db   : : : : :
Qy 561 GAG---GAGATCAACTGGAGCCATTACAGAGCGTGAGCCCGCGCGGAGGAGAACTTG 617
Qy 404 Thr-----ThrGluAsnArgPheHisSerLeu-ProPheSerLeuTh 417
Db   : : : : :
Qy 618 ACGTGGAGCAGCGCGCGGACGAGAGAGGTGCTCCCTTCAATCCCTTCGCTGTCTAC 677
Qy 417 rLysMetProAsn-----ThrAsnGlySerIleGlyHisSerProLeuSe 432
Db   : : : : :
Qy 678 AGCAGCTCTCCCGCTTTCGCGCGCCGCAAGCCCGCTCGCGCCCGCGCGCGCG 737
Qy 432 rLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProPr 452
Db   : : : : :
Qy 738 CGCTCCCGCAGCAGCGCTGGGCTCTCGGCGCCACCC-CCGCTTCAGCGCGCGCCCG 796
Qy 452 oLeuAlaMetProPro-----GlyAsnSerHi 461
Db   : : : : :
Qy 797 CCGCGCGCCACCCCGCCCGCCACCGCGCGCGCGCGCGCGCTGGGTGCGCGCG 856
Qy 461 sGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValI 481
Db   : : : : :
Qy 857 CAGATCCCGCGCGCTGGCTCCAGCGCTCAGACCGGAGAGC-----TGCTCTGGTGACCT 910
Qy 481 eArgTrpIleGlyGlnProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAs 501
Db   : : : : :
Qy 911 AGACGGTCCGGGGATCTCGCGCGCTTAGGGACTGTTGCTGGAGGTGAGTTGGTCA 970
Qy 501 pGluCysAlaGlyCysThr----- 507
Db   : : : : :
Qy 971 GGGTCCCACAGCGTCTCTCATGTGGAGAGCTTTAAAGTAGGTAAGAACTGGCAGAAGAA 1030
Qy 508 -----AspGlyThr-----PheArgG1 513
Db   : : : : :
Qy 1031 CCTGAGGTTTCATCTACAGCGTTTCTGTTGGAGTGGGACCCCGCAGAGACTAGGAAACGTAA 1090
Qy 513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCys-A 533
Db   : : : : :
Qy 1091 AGCAAAGTCATGTCATCTGTACATGTGTAGTACCCATATGAAACAGACTCCACTCTGTCT 1150
Qy 533 rGProAspSerArgPhe---AlaSerLeuGlnProValSerAsnGlnIleGluArgCysA 552
Db   : : : : :
Qy 1151 CTCCTGTGCTTTTGGCTGCTTCTACTGAGAAACATATTTCACAAACATGCAGAAACAAA 1210
Qy 552 snSer-----LeuAlaPheGlyTyrLeuSerGluValValGluG 566
Db   : : : : :
Qy 1211 GCAGCACCACTTATGCTGTAGACCTTTTTCATGCGGGTCATATATTGCTTTCATGTGAAGGA 1270
Qy 566 lu-----AsnThrProProLysMetGluLys-LeuGlyLeuGlu 578
Db   : : : : :
Qy 1271 TTATGTATATGCAAAAGACATAGACAGATTGCGCAAGAAACAAAGAAAAATTTTGAG 1330
Qy 579 IleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSer 598
Db   : : : : :
Qy 1331 ATT-----ATTAACTTCCACTCAGAGATGTTTC---TCATCA 1366
Qy 599 ThrLeuPheCysLeu-----PheAlaPhe 606
Db   : : : : :
Qy 1367 ACAGTTTATGACATCAGGTTTGAAGACAAAGCAATCAACCTGTGAGACAAAGGAACAGGA 1426
Qy 607 SerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyr 626
Db   : : : : :
Qy 1427 GCCAAATTTGGTGAAACC-----CAAGAAAAAGAGAAAAAGTTCAGTCTATAC 1477
Qy 627 -----SerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640
Db   : : : : :
Qy 1478 TTAGGCCCTGAGAGGGCTAATCAATCTTGGAAACACTTGTATTATGAATTTGATGTCCTCA 1537
Qy 641 LeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeu-----ArgLys 657
Db   : : : : :
Qy 1538 ---GGCACTTACCATTCTCTCTACTGAAGATTTCTTCTCTCTGACAGCACAATG 1594
Qy 658 IleLeuGluLysValGluAlaAlaSerGlyPhe----- 668
Db   : : : : :
Qy 1595 TATAATGACAAGCCCGAGCTTGTCTGTGCTGTGAAATGCTTCGCTTTTTCATGCTAT 1654
Qy 669 -----ThrSerGluGluLysAspPro 675
Db   : : : : :
Qy 1655 GTACTCTGGAGCGCAACTCTCTCACATTCCTATAAGTTACTGCATCTGATATGATGTCCTCA 1714
Qy 676 -GluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysI 695
Db   : : : : :
Qy 1715 TGCAGAACATT----- 1727
Qy 695 eArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAs 715
Db   : : : : :
Qy 1728 -----GCAGGCTACAGGCGAGGATGCCATGAGTCTCTTATTGCAATA----- 1772
Qy 715 nGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAs 735
Db   : : : : :
Qy 1773 -----TTAGCGTGTACATAGACAGCAAGATGATAGTGG 1810
Qy 735 nLeuLysPheAlaGluAlaPro-----SerCysLeuIleIleGlnMetProArgph 752
Db   : : : : :
Qy 1811 TGGCAGGAGGCCAATAAACCCCACTGCTTAAGTCTGATCATAGACCAATCTTTACAGG 1870
Qy 752 eGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuLeuLeuAsnIleThr-- 771
Db   : : : : :
Qy 1871 TGGCTGCAATCAGATGTCACATGTCAGCGCTGCCATAGTGTCTTCTACCACCATAGACCC 1930
Qy 772 -----AspLeuLeuGluAspThrProArgGlnCys----- 781
Db   : : : : :
Qy 1931 ATGCTGGGACATCAGTTTGACTTGGCTTGGCTGCTCTGTGCCACATTCGATTTCCCAAGACC 1990
Qy 782 -----ArgIleCysGlyGlyLeuAlaMetTy 790
Db   : : : : :
Qy 1991 AGAGAGGCTGACAGCACAGTGCAGAGGAGTGCACCATACAGGAATCCCTCCTTACCTTAC 2050
Qy 790 rGluCysArgGluCysTyrAspAspProAspIle-----SerAlaGlyLysIleLysG1 808
Db   : : : : :
Qy 2051 AGACTGTCTACAGTGGTTTACAGGCCAGACACCTAGGAGCAGTGCCTCAAAATCAAA-- 2108
Qy 808 nPheCysLysThrCysAsnThrGlnValHis----- 818
Db   : : : : :
Qy 2109 ---TGAATAGTTGCCAAAGCTACCAAGAGTCTTACTAAACAGCTCACAATGAAAAAATT 2164
Qy 819 -----LeuHisProLysArgLeuAsnHis-----Ly 827
Db   : : : : :
Qy 2165 ACCATTGTGGCTGTTTTCATCTCAAGCGGTTTGAGCATGTAGGCACAGAGCGGAA 2224
Qy 827 sTyrAsnPro---ValSerLeuProLysAspLeu----- 837
Db   : : : : :
Qy 2225 GATTATACCTTTATCTCTTTCCTTGGAGCTGCACATGACTCCGTTTTTGGCCTCTTAC 2284
Qy 838 -----ProAspTrpAspTrpArgHisGlyCysIlePr 848
Db   : : : : :
Qy 2285 TAAAGAGAGCAGAATGAAGAGGCCAGCCACCAACAGAT-----TGTGTGCC 2332
Qy 848 oCysGlnAsn---MetGluLeuPheAlaValLeu-----CysIleGluThrSe 863
Db   : : : : :
Qy 2333 CAATGAGAAAGATTTCTCTTGTGAGTATTAATCACCATGGAACCTTGGAAAGTGG 2392
Qy 863 rHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPhePheAsp 882
```


QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnPro 830
 Db 1408 ACACATCTACTCCCAAGAGAAAATAGTATGATGGTGAATCCA 1449

RESULT 11

US-09-262-537-57
 ; Sequence 57, Application US/09262537
 ; Patent No. 6479256
 ; GENERAL INFORMATION:
 ; APPLICANT: Hayflick, Joel
 ; TITLE OF INVENTION: Lectomedin Materials and Methods
 ; FILE REFERENCE: 27866/35307
 ; CURRENT APPLICATION NUMBER: US/09/262,537
 ; CURRENT FILING DATE: 1999-03-04
 ; EARLIER APPLICATION NUMBER: 60/076,782
 ; EARLIER FILING DATE: 1998-03-04
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 57
 ; LENGTH: 5610
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (281)...(4687)
 US-09-262-537-57

Alignment Scores:
 Pred. No.: 0.0135 Length: 5610
 Score: 128.50 Matches: 139
 Percent Similarity: 34.51% Conservative: 86
 Best Local Similarity: 21.32% Mismatches: 224
 Query Match: 2.55% Indels: 204
 DB: 4 Gaps: 37

US-09-671-687A-3 (1-949) x US-09-262-537-57 (1-5610)

QY 172 GlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAsp 191
 Db 902 GGCACAGGCTTTGTG-----GTCTACGATGGTGGCCGCTTGTACAAACAGGAGCGCAGC 955
 QY 192 CysGlyPheValAlaLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeuGlu 211
 Db 956 CGCAACATCGTCAAGTATGAC----- 976
 QY 212 SerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuProProLeuGluIleAsn 231
 Db 976 ----- 976
 QY 232 SerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePheCysAsp 251
 Db 977 CTACGACGCGCATCAAGACGGGGAGACGGTC---ATCAATACCGCAACTACCATGAC 1033
 QY 252 ValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIle 271
 Db 1034 ACCTGCCCTACCGC----- 1048
 QY 272 GlyAsnTrpAspGlyArgPheAsp-----GlyValLeuCys 283
 Db 1049 -----TGGGGCGAAGACCGACATTGACCTGGCGGTGACGAGACGGGCTGTGGGTC 1102
 QY 284 SerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSer 303
 Db 1103 ATCTACGCCACTGAGGCAACAACGGCGGCTGTGGTGTGAGCCAGCTGAACCCCTACACA 1162
 QY 304 ValThrGlnGlu-----ArgArgProProLysLeuAlaPhe 315
 Db 1163 CTGGCTTTGGGGCAGCTGGGAGACGGGTTACGACAAAGCGCTCGGCATCAACGCCCTTC 1222
 QY 316 MetSerArgGly-----ValGlyAspLysGlySerSer 326
 Db 1223 ATGGTGTGGGGTCTGTGTAGCTCTCGTTCGTGTGATGGTATGATGACAGCGAGCGC 1282

QY 327 SerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArgSerGlu 346
 Db 1283 GCTGGCAACCGGTGGACTATGCTTCAACACCAAT-----GCCAACCGGAGGACGCT 1336
 QY 347 LeuPheTyrThrLeuAsnGly-----SerSerValAspSerGlnProGln 361
 Db 1337 GTCAGCCTCACCTTCCCAACACCCCTACCAAGTTTCATCTCTCTCTGACTACTAAACCTTCGC 1396
 QY 362 SerLysSerLysAsnThrTrp-----TyrIleAspGluValAlaGluAspProAlaLys 379
 Db 1397 GACAACACAGCTGACGCTGGAAACAATATTC----- 1429
 QY 380 SerLeuThrGluIleSerThrAspPhe-----AspArgSerSerProProLeuGln 396
 Db 1430 ---GTGGTGGCTACAGCTGGAGTTCCGGCCGCCGCCAGTGTGCGCCAGCCACT 1486
 QY 397 ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 Db 1487 TCCCAACCCCTCAGACGACACCAACAGCC-----AGGCCACGCGCTCCACGAGC 1537
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 Db 1538 ACAGCTCGCCCGCAGCCACCCCGCTCCGGCGGACCCCTCACCACGACCCAGTG 1597
 QY 437 SerValMetGluGluLeuAsn-----ThrAlaProValGlnGluSer 450
 Db 1598 GGTGCCATCAACACAGCTGGGACCTGATCTGCTCCAGCCACAGCCCGCTC----- 1648
 QY 451 ProProLeuAlaMetProProGlyAsnSer-HisGlyLeuValGlySerLeuAlaGln 470
 Db 1649 CCCAGCACCGCGGGCCCCCGACCGAATCTACAGTGTCCCTGAGCTCTTCGCG-- 1706
 QY 470 uValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLe 490
 Db 1707 -----AGCCCGGAGAGGTACGCGG-GTCCAGTGG-----CCGGCCAC 1743
 QY 490 uAsnGluValLeuAlaGlyLeuGluAspGluCysAlaGlyCysThrAspGlyTrp 510
 Db 1744 CAGCAG-----GGCATGTGTGGAG-----AGGCCCTGCCCAAGGGAC 1785
 QY 510 rPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPhe----- 526
 Db 1786 T---CGAGGAATTCCTCTCCAGTGTCTACACAGCTTGGGGCTCTGGAAACCCCGGG 1842
 QY 527 -ValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAs 546
 Db 1843 CCTGACCTCAGCAACTGC-----ACCTCCCTCCCTGGTCAA 1878
 QY 546 nGlnIle-----GluArgCysAsnSerLeuAla----- 555
 Db 1879 CCAGTGGCCCAAGATCAAGAGTGGGGAGAACCGCGGCCAACATCCCGCAGAGTGGC 1938
 QY 556 -----PheGlyGlyTyrLeuSerGluValValGluGluAsnTh 568
 Db 1939 CCAGACACCCCGGGCTCCATCTACGGCGGGAGAGCTCTCTCTCTCTGTG----- 1987
 QY 568 rProProLysMetGluLysGlyGluLeuGluIleMetIleGlyLysLysGlyIleGln 588
 Db 1988 -----AAGCTGATGAGCAGCTGTGTGGACATCTCTGGATGCCAGCTGACGGCCCTGCG 2040
 QY 588 n-----GlyHisTyrAsnSerCysTyrLeuAspSerThrLe 600
 Db 2041 GCCCATCGCGGAGTACCGCGCAAGAACTACACAAGATGCACAGCGCAGGAGAGAAC 2100
 QY 600 uPheCysLeuPheAlaPheSerSerValLeuAspThrVal-----LeuLeuArgProLys 618
 Db 2101 T---TGTAAGGATTATATCAAGCGCGGTGGAGACAGTGGACAATCTGTCTCCGCCCA-- 2155
 QY 618 sGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleVal 638
 Db 2156 -----GAAGCTCTGGAGTCTCTGGAGGACATGAATGCCACGGAGCAGGT 2199


```
QY 638 lAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysI 658
Db 2200 GCAC-----ACGGCCACCAGTCTCTC-----GACGT 2226
QY 658 eLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPh 678
Db 2227 CTGGAG-----GAGGGCGCTTCTGCTGGCGCAATGTCTAGGAGCCTCCGCTT 2280
QY 678 eLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAl 698
Db 2281 CCTGGCTGCAAGGAGACGCTGCTGGAGGTACAGTCTCTG-----AACACAGA 2331
QY 698 aGlyGlnLysValGlnAspCysTyrPheTyrGln-----IlePheMetGluLysAsnGlu-- 716
Db 2332 GGGCCAG--GTGCAGGAGCTGTGTCTCCCGAGGAGGTACCCCGAGAAAGAACTCCAT 2388
QY 717 -LysValGlyValProThrIleGlnGln-----LeuLeuGluTyrSerPheIleAsnS 734
Db 2389 CAGCTGTCTGCCAAACCATCAAGCAGACAGCCGCGCATGGGTGGTCAAGTTGTCTT 2448
QY 734 eAsnLeuLysPheAlaGluAlaProSerCys 744
Db 2449 CATCTCTACAAACACCTGGCGCTCTCTCTGT 2480

RESULT 12
US-08-999-774A-11
; Sequence 11, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: McLanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; NUMBER OF INVENTIONS: Related Reagents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,774A
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,818
; FILING DATE: 11-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-9196
; TELEFAX: (650) 496-1200
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1503 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 25..1503
```

```
US-08-999-774A-11
Alignment Scores:
Pred. No.: 0.00228 Length: 1503
Score: 126.00 Matches: 77
Percent Similarity: 34.86% Conservative: 37
Best Local Similarity: 23.55% Mismatches: 84
Query Match: 2.50% Indels: 129
DB: 3 Gaps: 21

US-09-671-687A-3 (1-949) x US-08-999-774A-11 (1-1503)
QY 255 GlyLysGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIle---GlyAsn 273
Db 118 GGCTCGATTCTCAGGTTATTAT-----GACCAGGAATAATTTATGGTGA 162
QY 274 TrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu----- 289
Db 163 AGTGACAGCAGATTGCTGGATAGTGACATCAATTCTGCACTGAACCTTGAAGATGAT 222
QY 290 -----SerThrIleLeuLeu-----His--- 295
Db 223 GACGATGACTATTTCATCTACGAGTTTCTGGTCAGAAAGACCAGGATATCATGCC 282
QY 296 -----IleAsnAspIleIleProGluSerValThrGln----- 306
Db 283 CCTGTGGCATTGCTTAATGAT---ATACCAGGTCAACAGAACAGTATGATCCTTTGCT 339
QY 307 GluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSer 326
Db 340 GAGCACAGACCTCCAAAGATTGCA-----GACCGG---GAAGAT 375
QY 327 SerHisAsnLysProLysAlaThr----- 334
Db 376 GAATACAAAAGCATAGGCGGACCATGATAATTTCCCAGAGCGTCTTGATCCTTTTGA 435
QY 335 -----GlySerThrSerAspProGly---AsnArgArgSerGluLeu----- 347
Db 436 GATGGAGGGAAGACCCCTGATCCTAAATGAATGCTAGGACTTACATGGATGTTATGCGA 495
QY 347 ----- 347
Db 496 GAACAACACTTGACTAAAGAAACGAGAAATTAGGCAACAGCTAGCAGAAAAAGCTAAA 555
QY 348 -----PheTyrThrLeuAsnGlySerValAspSerGlnProGlnSerLysSer 364
Db 556 GCTGGGAGAACTAAAAGTCGTCATGGAGCAGCAGCG---TCCCAGCTCCATCAAAACGA 612
QY 365 LysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIle 384
Db 613 AAACGGCGTTGG-----GATCAACAGCTGATCAGACTCCTGTCCTCCCAAAAAA 666
QY 385 SerThrAspPheAspArgSerSerProLeuGlnProProProValAsnSerLeuThr 404
Db 667 CTATCAAGTTGGGATCAGCAGACCCCTGGGCATACCTCTCTTAAGATGGGATGAG 726
QY 405 ThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGly 424
Db 727 ACACCGGTCGT-----GCAAGGGAAGGAGACTCCTGGA 762
QY 425 SerIle-----GlyHisSerProLeuSerLeuSer 434
Db 763 GCAACCCAGGCTCAAAAATATGGGATCCTACACCTAGCCACACACCCAGCGGAGCTGCT 822
QY 435 AlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAla 454
Db 823 -----ACTCTGGACGAGGTGATACACAGGCCAT 852
QY 455 MetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsn 474
Db 853 GCGACACCCCGGC-----CATGGA-----GGCGCAACTTCCAGTGTCTGTAATAAAC 897
QY 475 ProProPheTyrGlyValIleArgTrp-IleGlyGlnProProGlyLeuAsnGluValLe 494
```

```
Db 898 -----AGATGGATG---AAACCCCAAAACAGAGAGATACT 933
QY 494 uLaGlyLeuGluLeuGlu 500
Db 934 CTTGGCATGGAAGTGGAT 952

RESULT 13
US-09-976-594-799
; Sequence 799, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 799
; LENGTH: 4282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1385527.4
US-09-976-594-799

Alignment Scores:
Pred. No.: 0.0157 Length: 4282
Score: 126.00 Matches: 77
Percent Similarity: 34.86% Conservative: 37
Best Local Similarity: 23.55% Mismatches: 84
Query Match: 2.50% Indels: 129
DB: 4 Gaps: 21

US-09-671-687A-3 (1-949) x US-09-976-594-799 (1-4282)
QY 255 GlyLeuSerLeuGlyTyrPheValGlyValAspMetAspAsnProLeu---GlyAsn 273
Db 126 GGCTCGATTCTCAGGTTATTAT-----GACCAGGAATTTATGTGGA 170
QY 274 TrpAspGlyArgPheAspGlyValLeuGlySerPheAlaCysValGlu----- 289
Db 171 AGTGACAGCAGATTGCTGGATACGTGACATCAATCTGCTGCAACTGAACTTGAAGATGAT 230
QY 290 -----SerThrIleLeuLeu-----His--- 295
Db 231 GACGATGACTATTCATCATCTACGAGTTCTGCTGGTCAGAAAGCCAGGATATCATGCC 290
QY 296 -----IleAsnAspIleIleProGluSerValThrGln----- 306
Db 291 CCTGTGGCATTGCTTAATGAT---ATACCACAGTCACACAAAGATGATCATCTTGCT 347
QY 307 GluArgArgProGlyValLeuAlaPheMetSerArgGlyValGlyAspGlySerSer 326
Db 348 GAGCAGACAGCTCCAAAGATTGCA-----GACCGG---GAAGAT 383
QY 327 SerHisAsnLysProLysAlaThr----- 334
Db 384 GAATACAAAAGCATGATGGCGGACCATGATAATTTCCCGAGCGCTTTGATCCTTTTGGCA 443
QY 335 -----GlySerThrSerAspProGly---AsnArgArgSerGluLeu----- 347
Db 444 GATGGAGGGAAGACCCCTGATCTTAAATGAATGCTAGGACTTACATGGATGTAATCGGA 503
QY 347 ----- 347
Db 504 GAACAACACTTGACTAAAGAAAGACGAGAAATTAGGCAACAGCTAGCAGAAAAGCTAAA 563
```

```
QY 348 -----PheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSer 364
Db 564 GCTGGAGAACTAAAGTCGTCAATGAGCGACGCG---TCCAGCGCTCCATCAAAACGA 620
QY 365 LysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIle 384
Db 621 AAACGGCGTTGG-----GATCAACAGCTGATCAGACTCCTGCTGCGCACTCCCAAAAA 674
QY 385 SerThrAspPheAspArgSerSerProProLeuGlnProProValAsnSerLeuThr 404
Db 675 CTATCAAGTTGGGATCAGGACAGAGACCCCTGGGCATACTCTTCTTAAGATGGGATGAG 734
QY 405 ThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGly 424
Db 735 ACACCAAGGTCGT-----GCAAAAGGGAAGCAGAGACTCTCTTGA 770
QY 425 SerIle-----GlyHisSerProLeuSerLeuSer 434
Db 771 GCAACCCCGAGGCTCAAAAATATGGGATCTTACACTAGCCACACACCGCGGAGCTGCT 830
QY 435 AlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProProLeuAla 454
Db 831 -----ACTCTGGAGCAGGTGATACACAGGCCAT 860
QY 455 MetProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsn 474
Db 861 GCGACACACAGCG-----CATGGA-----GGCGCAACTTCCAGTGTCTGTAATAAAC 905
QY 475 ProProPheTyrGlyValIleArgTyrPheGlyGlnProProGlyLeuAsnGluVal 494
Db 906 -----AGATGGATG---AAACCCCAAAACAGAGAGATACT 941

RESULT 14
US-09-949-016-1717
; Sequence 1717, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1717
; LENGTH: 6474
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1717

Alignment Scores:
Pred. No.: 0.0337 Length: 6474
Score: 126.00 Matches: 77
Percent Similarity: 34.86% Conservative: 37
Best Local Similarity: 23.55% Mismatches: 84
Query Match: 2.50% Indels: 129
DB: 4 Gaps: 21

US-09-671-687A-3 (1-949) x US-09-949-016-1717 (1-6474)
QY 255 GlyLeuSerLeuGlyTyrPheValGlyValAspMetAspAsnProLeu---GlyAsn 273
```

```
Db 134 GGCTCGATTCTACAGTTATTAT-----GACCAGGAAATTTATGTGGA 178
Qy 274 TrpAspGlyArgPheAspGlyValLeuCysSerPheAlaCysValGlu----- 289
Db 179 AGTGACAGAGATTTGCTGATACGTCAGATCAATGCTGCAACTTGAAGATGAT 238
Qy 290 -----SerThrIleLeuLeu-----His----- 295
Db 239 GACGATGACTATTCATCACTACAGTTGCTTGGTCAGAAAGCCAGGATATCATGCC 298
Qy 296 -----IleAsnAspIleIleProGluSerValThrGln----- 306
Db 299 CCTGTGGCATTGCTTAATGAT---ATACACAGCTCAACAGACAGTATGATCCATTGCT 355
Qy 307 GluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSer 326
Db 356 GAGCAGACAGCTCTCAAGATTGCA-----GACCGG---GAAGAT 391
Qy 327 SerHisAsnLysProLysAlaThr----- 334
Db 392 GAATACAAAGATAGCGGACCATGATAATTTCCCGAGAGCGTCTTGATCCTTTTGCA 451
Qy 335 -----GlySerThrSerAspProGly---AsnArgArgSerGluLeu----- 347
Db 452 GATGGAGGAAACCCCTGATCTTAAATCAATGCTAGGACTTACATGGATGTAATGCGA 511
Qy 347 ----- 347
Db 512 GAACACACTTGACTAAAGAAAGACAGAAATTTAGGCAACAGCTAGCAGAAAAAGCTAA 571
Qy 348 -----PheTyThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSer 364
Db 572 GCTGGAGAACTAAAGTCTGTCATGGAGCAGCAGCG---TCCGAGCGCTCCATCAAAACGA 628
Qy 365 LysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIle 384
Db 629 AAACGGCGTTGG-----GATCAACAGCTGATCAGACTCCTGTCGCCACTCCCAAAAA 682
Qy 385 SerThrAspPheAspArgSerSerProProLeuGlnProProProValAsnSerLeuThr 404
Db 683 CTATCAAGTTGGATCAGCAGAGACCCCTGGGCATACTCTTCTTAAAGTGGATGAG 742
Qy 405 ThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGly 424
Db 743 ACACCAAGGTCGT-----GCAAGGGAAGCGAGACTCCTCTGGA 778
Qy 425 SerIle-----GlyHisSerProLeuSerLeuSer 434
Db 779 GCAACCCCGCTCAAAATATGGGATCTTACACCTAGCCACACACAGCGGGAGCTGCT 838
Qy 435 AlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGlnSerProProLeuAla 454
Db 839 -----ACTCTGGAGCGGTTGATACACACCGCCAT 868
Qy 455 MetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsn 474
Db 869 GCGACACCAAGGC-----CATGCA-----GGCGCAACTTCCAGTGTCTGTAATAAAC 913
Qy 475 ProProPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValle 494
Db 914 -----AGATGGGATG---AAACCCCAAAACAGAGAGAGATACT 949
Qy 494 uAlaGlyLeuGluLeuGlu 500
Db 950 CCTGGCGATGGAAGTGGAT 968
```

RESULT 15

```
US-09-270-767-15085
; Sequence 15085, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
```

```
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15085
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15085
```

Alignment Scores:

```
Pred. No.: 0.00537 Length: 1463
Score: 122.50 Matches: 69
Percent Similarity: 38.87% Conservative: 34
Best Local Similarity: 26.04% Mismatches: 89
Query Match: 2.43% Indels: 73
DB: 4 Gaps: 16
```

US-09-671-687A-3 (1-949) x US-09-270-767-15085 (1-1463)

```
Qy 229 GluIleAsnSerArgValSerLeuLysGlyGluThrIleGluSerGlyThrValIle 248
Db 568 GAAATCTGTCCAGCGTG-----GGCAAGGAGAACACCCAGCCCAACGCCGTT 618
Qy 249 PheCysAspValLeuProGlyLysGluSerLeu----- 259
Db 619 TCGGCCACACGACGCTGGCAAAAAAATCTCATCGAACTCATCTCGGCAGGTGG 678
Qy 260 GlyTyThrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAsp 279
Db 679 TCGCACTTT---GGAGCAGCGTCCGACAAAGAACACCAACCCACCCAGCAGTACGAC 735
Qy 280 -GlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAsp-- 298
Db 736 ATCG-----GGTCGGTTGCGCTGTGATACCAACGACAA 768
Qy 299 -----IleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMe 316
Db 769 TCATGTAAAAAACACCAAGCACCACCAAGTATGCCATGCTCCGCCAGAG----- 817
Qy 316 tSerArgGlyValGlyAspLysGlySerSerSerHis---AsnLysProLysAlaThrG 335
Db 818 -TCCACGCCACATCCCCACCGGTCAGCAGATATGTAATGGAAATGCTCAGAATGG 876
Qy 335 ySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyThrLeuAsnGlySerSe 355
Db 877 AGTACACGCGGA-----AATGGAAGTGG 900
Qy 355 rValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyIleAspGluValAlaG 375
Db 901 CCAGGCGCGCAGCCCCCAGAGTGTCTAGT---AACGGATGACGCAAGAGATCGCGCAAC 957
Qy 375 uAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProPro-- 394
Db 958 GACTACTTTCACCTTCAACTACA-----ACCATCTCTCCAG 993
Qy 395 ----LeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPr 413
Db 994 GCATAGTCAAGTCCGCCCACTCCAATATA-----TCGCGGCC 1032
Qy 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerL 433
Db 1033 AGCTTCAGTGACA-----CCCTCTGCAATGGAATGTG---CACAGTCCGAATGCGAA 1083
Qy 433 euSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProProL 453
Db 1084 CAGCACACCGCGGGAAGCGGAGAG---GAAGTAGATCTCTATACAGCAGCAACGCCCG 1140
Qy 453 euAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysG 473
Db 1141 TC-----AACGCTCTCAGCGTCCGAGCTCGCGGCAACAAGGTGA 1179
```

QY 473 luAsnProPro 476
: : : : :
Db 1180 GGAGICTTCA 1190

Search completed: April 18, 2005, 20:13:32
Job time : 425 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:14:54 ; Search time 73 Seconds
(without alignments)
5027.885 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQEKVTSPWEERI.....RLLCDAYMCYQSPWLSLYK 949

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq19808:*
- 2: Geneseq19908:*
- 3: Geneseq20008:*
- 4: Geneseq20018:*
- 5: Geneseq20028:*
- 6: Geneseq20038:*
- 7: Geneseq20038s:*
- 8: Geneseq20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4983	99.0	953	7	ADC24816
2	4983	99.0	953	8	ADQ95918 Human bre
3	4983	99.0	960	8	ADR99244 DKFZp586D
4	4980	98.9	953	8	ADQ95916 T cell ac
5	4971.5	98.8	956	6	ABB82783 Human CYL
6	4971.5	98.8	956	8	ADRI4489 Human NF-
7	4968.5	97.7	956	8	ADQ95920 T cell ac
8	3876.5	77.0	739	4	AU23747 Novel hum
9	3862	76.7	731	3	AAY91954 Human cyt
10	3855	76.6	731	4	AAB95828 Human pro
11	3636	72.2	685	4	AAM39234 Human pol
12	3563	70.8	698	4	AAM41040 Human pol
13	2755	54.7	558	4	AU23211 Novel hum
14	2343	46.5	476	4	AAB95719 Human pro
15	1275	25.3	261	5	ABB89233 Human pol
16	1173	23.3	238	5	ABB89234 Human pol
17	856	17.0	517	4	ABB61669 Drosophil
18	597	11.9	113	4	AAM14965 Peptide #
19	597	11.9	113	4	ABB33941 Peptide #
20	597	11.9	113	4	AAM27399 Peptide #
21	597	11.9	113	4	ABB28754 Peptide #
22	597	11.9	113	4	ABB19377 Protein #
23	597	11.9	113	4	AAM67104 Human bon
24	597	11.9	113	4	AAM54704 Human bra
25	597	11.9	113	4	AAM02691 Peptide #

26	597	11.9	113	5	ABG36764	Abg36764 Human pep
27	568	11.3	106	8	ADK71947	Adk71947 Human ori
28	545	10.8	101	8	ADK71945	Adk71945 Human ori
29	514	10.2	96	8	ADK71949	Adk71949 Human ori
30	492	9.8	104	8	ADK71962	Adk71962 Human ori
31	491	9.8	91	8	ADK71943	Adk71943 Human ori
32	151	3.0	1392	2	AAY06999	Aay06999 Restin pr
33	151	3.0	1427	8	ADP56353	Adp56353 Human PRO
34	149	3.0	1427	2	AAR10534	Aar10534 Human 160
35	146	2.9	354	7	ABM85235	Abm85235 Human pro
36	143	2.8	547	5	AAU74342	Aau74342 Human cyt
37	143	2.8	547	5	ABB97353	Abb97353 Novel hum
38	143	2.8	547	7	ADM04066	Adm04066 Human pro
39	139.5	2.8	1921	4	ABB62962	Abb62962 Drosophil
40	136	2.7	708	8	ADN99865	Adn99865 Novel hum
41	135.5	2.7	306	3	AAB42642	Aab42642 Human ORF
42	135.5	2.7	306	7	ADB64836	Adb64836 Human pro
43	135.5	2.7	721	4	AAU87339	Aau87339 Novel cen
44	135.5	2.7	721	8	ADI54654	Adi54654 Novel hum
45	135.5	2.7	2273	6	ABU38191	Abu38191 Protein e

ALIGNMENTS

RESULT 1
ADC24816

ID ADC24816 standard; protein; 953 AA.

XX AC ADC24816;

XX DT 18-DEC-2003 (first entry)

XX DE Human breast specific polypeptide (BSP) DEX0238_137, SEQ ID NO:137.

XX KW Human; breast specific polypeptide; BSP; breast specific nucleic acid;
 KW BSNAs; breast cancer; metastasis; non-cancerous disease; breast tissue;
 KW identification; monitoring; diagnosis;
 KW engineered breast tissue production; transgenic animal; drug screening;
 KW cytostatic; gene therapy; vaccine; chromosome 16p13.3.

XX OS Homo sapiens.

XX PN WO2003020900-A2.

XX PD 13-MAR-2003.

XX PF 29-AUG-2002; 2002WO-US027777.

XX 31-AUG-2001; 2001US-0316306P. - *late*

XX (DIAD-) DIADEXUS INC.

XX PI Sun Y, Liu C, Salceda S;

XX WPI; 2003-290182/28.

XX N-PSDB; ADC24898.

XX PT New breast specific polypeptide useful for identifying, diagnosing,
 PT monitoring, staging, imaging and treating breast cancer and non-cancerous
 PT disease states in breast.

XX PS Disclosure; SEQ ID NO 137; 264pp; English.

XX CC The invention relates to breast specific polypeptides (BSPs) and nucleic
 CC acids (breast specific nucleic acids; BSNAs) encoding them. The invention
 CC also relates to vectors and host cells comprising a BSNAs sequence;
 CC antibodies against BSPs; the recombinant production of BSPs; methods of
 CC detection of BSNAs or BSPs in a sample; kits for detecting a risk of
 CC cancer or presence of cancer in a patient; and vaccines comprising a BSNAs
 CC or BSP. The invention additionally discloses fragments, mutants, fusion
 CC proteins, homologous proteins and allelic variants of BSPs; methods for
 CC identifying and designing agonists and antagonists of BSPs; methods for

QY 1 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKQTKLLKVPKGSIGQYIDRSVGHRI 60
Db 1 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKQTKLLKVPKGSIGQYIDRSVGHRI 60
QY 61 PSAGKKNQIGLKIQLQPHAVLFVDE - DVVINEKFTTELLAITNCEERFSLFKNNRLS 119
Db 61 PSAGKKNQIGLKIQLQPHAVLFVDEKDVVINEKFTTELLAITNCEERFSLFKNNRLS 120
QY 120 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVELLEEGRGQFTDGV 179
Db 121 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVELLEEGRGQFTDGV 180
QY 180 YQKQLFOCDEDCG - FVALDKLELIEDDDTALESYAGPGDTMQVELPLEINSRVSLKG 238
Db 181 YQKQLFOCDEDCGVFVALDKLELIEDDDTALESYAGPGDTMQVELPLEINSRVSLKV 240
QY 239 GETIESGTIVFCVLPKGSIGVFGVDMNDPNTGNWDRFDGV - LCSFACVESTILLHN 297
Db 241 GETIESGTIVFCVLPKGSIGVFGVDMNDPNTGNWDRFDGV - LCSFACVESTILLHN 300
QY 298 DIIPESVTQERRPPKLAFLMSRGVGDGSSSHNPKATGSTDPGNNR - RSELFYTLNGSSV 356
Db 301 DIIPESVTQERRPPKLAFLMSRGVGDGSSSHNPKATGSTDPGNNRSELFTYTLNGSSV 360
QY 357 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPIOPPVPVNSLTENRPHSLPFSL 416
Db 361 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPIOPPVPVNSLTENRPHSLPFSL 420
QY 417 TQMPNTNGSIGHSPLSLSAQSWMEELNAPQESPLAMPNGSHGLEVGS LAEVKENP 476
Db 421 TQMPNTNGSIGHSPLSLSAQSWMEELNAPQESPLAMPNGSHGLEVGS LAEVKENP 480
QY 477 FYGVIRWTGQPLNEVLAGLEBECACGTGTGTRTYFTCALKKALFVKLKS CRPDS 536
Db 481 FYGVIRWTGQPLNEVLAGLEBECACGTGTGTRTYFTCALKKALFVKLKS CRPDS 540
QY 537 RFASLPQVSNQIERCNSLAFGGYLSVVEENTPPPKMEKEGLEIMICKKGIQGHYNSCYL 596
Db 541 RFASLPQVSNQIERCNSLAFGGYLSVVEENTPPPKMEKEGLEIMICKKGIQGHYNSCYL 600
QY 597 DSTLFCFAFSSVLTLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVVCATKIMKLR 656
Db 601 DSTLFCFAFSSVLTLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVVCATKIMKLR 660
QY 657 KILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAGQKQDCVYQIFMEKNE 716
Db 661 KILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAGQKQDCVYQIFMEKNE 720
QY 717 KVGVTPIQOLLEWSFINSNLKFAEAPSCLI IQMPFRGKDFKLPKIFPSSLELNIITDLED 776
Db 721 KVGVTPIQOLLEWSFINSNLKFAEAPSCLI IQMPFRGKDFKLPKIFPSSLELNIITDLED 780
QY 777 TPRQCRICGLAMYECRECYDDPDTSAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVS LPKD 836
Db 781 TPRQCRICGLAMYECRECYDDPDTSAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVS LPKD 840
QY 837 LPDWDWRHGCIPCONMELFVLCIETSHVYVAFVKYKDDSAWLPFDSMADRDGQNGFNI 896
Db 841 LPDWDWRHGCIPCONMELFVLCIETSHVYVAFVKYKDDSAWLPFDSMADRDGQNGFNI 900
QY 897 PQTTCPEVGEYKWSLEDLHSLDSRRIQGCARRLLCDAYMCYQPTMSLYK 949
Db 901 PQTTCPEVGEYKWSLEDLHSLDSRRIQGCARRLLCDAYMCYQPTMSLYK 953
RESULT 3
ADR99244
ID ADR99244 standard; protein; 960 AA.
XX
AC ADR99244;
XX
DT 02-DEC-2004 (first entry)

XX DKFP586D1122, SEQ ID 250.
DE
XX
KW Cytostatic; breast cancer; cancer; human; DKFP586D1122.
XX
OS Homo sapiens.
XX
PN WO2004078035-A2.
XX
PD 16-SEP-2004.
XX
PF 27-FEB-2004; 2004WO-US007268.
XX
PR 28-FEB-2003; 2003US-0450655P.
XX
PA (PARB) BAYER PHARM CORP.
XX
PI Eveleigh D, Bigwood D;
XX
XX WPI; 2004-653556/63.
DR N-PSDB; ADR99117.
DR
XX
XX Diagnosing breast cancer comprises comparing the level of expression of genes or gene products in a first biological sample taken from a patient with that in a normal patient sample.
XX
PS Claim 3; SEQ ID NO 250; 53pp; English.
XX
CC The present invention relates to a method (M1) for diagnosing breast cancer in a patient. The method comprises comparing the level of expression of one or more genes or gene products in a biological sample from the patient with that in a normal patient sample, where a difference in the gene expression in the first sample compared to that in the second sample is a diagnostic of the disease. Also claimed are: method (M2) for distinguishing between normal and disease tissues; method (M3) for monitoring the response of a breast cancer patient to treatment with an anti-cancer agent; method (M4) for identifying a compound for treating breast cancer; and an array for distinguishing between normal and disease tissues comprising two or more probes corresponding to genes selected from ADR98995-ADR99121 or comprising two or more polypeptides selected from ADR99122-ADR99248. In M1 and M2 the genes are selected from ADR98995-ADR99121 and the gene products are polypeptides selected from ADR99122-ADR99248. M1 is useful for diagnosing breast cancer. M2 and the array are useful for distinguishing between normal and disease tissue. M3 is useful for monitoring the response of a breast cancer patient to treatment with an anti-cancer agent. M4 is useful for identifying a compound for treating breast cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 960 AA;

Query Match 99.0%; Score 4983; DB 8; Length 960;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKQTKLLKVPKGSIGQYIDRSVGHRI 60
Db 8 MSSGLWSQEKVTSFYWEERIFVYLLLOECSTVDKQTKLLKVPKGSIGQYIDRSVGHRI 67
QY 61 PSAGKKNQIGLKIQLQPHAVLFVDE - DVVINEKFTTELLAITNCEERFSLFKNNRLS 119
Db 68 PSAGKKNQIGLKIQLQPHAVLFVDEKDVVINEKFTTELLAITNCEERFSLFKNNRLS 127
QY 120 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVELLEEGRGQFTDGV 179
Db 128 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVELLEEGRGQFTDGV 187
QY 180 YQKQLFOCDEDCG - FVALDKLELIEDDDTALESYAGPGDTMQVELPLEINSRVSLKG 238
Db 188 YQKQLFOCDEDCGVFVALDKLELIEDDDTALESYAGPGDTMQVELPLEINSRVSLKV 247
QY 239 GETIESGTIVFCVLPKGSIGVFGVDMNDPNTGNWDRFDGV - LCSFACVESTILLHN 297

Db 248 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDRFDGVLQCSFACVESTILLHN 307
Qy 298 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNR-RSELYFTLLNGSSV 356
Db 308 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNRSELFTLLNGSSV 367
Qy 357 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTFNRFHSLPFSL 416
Db 368 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTFNRFHSLPFSL 427
Qy 417 TKMPNTNGSIHSPUSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 476
Db 428 TKMPNTNGSIHSPUSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 487
Qy 477 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKKAFLVKLSCRPDS 536
Db 488 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKKAFLVKLSCRPDS 547
Qy 537 RFASLQPVSNQIERCNSLAFGGYLSSEVVEENTPPKMEKEGLEIMIGKKGIQHYNNSCYL 596
Db 548 RFASLQPVSNQIERCNSLAFGGYLSSEVVEENTPPKMEKEGLEIMIGKKGIQHYNNSCYL 607
Qy 597 DSTLFCFAFSSVLTLLRPKEKNDVEYYSQELLRTTEINPLRIYGVYCATKIMKLR 656
Db 608 DSTLFCFAFSSVLTLLRPKEKNDVEYYSQELLRTTEINPLRIYGVYCATKIMKLR 667
Qy 657 KILEKVEAASGTSREKDPPEELNITLPHILAVEPLKIRSAQKQVQDCVFYQIFWEKNE 716
Db 668 KILEKVEAASGTSREKDPPEELNITLPHILAVEPLKIRSAQKQVQDCVFYQIFWEKNE 727
Qy 717 KVGVTPIQQLLEWSEFINSNLKFAEAPSLIIQMPRGKDFKLFKIFPSLELNITDLED 776
Db 728 KVGVTPIQQLLEWSEFINSNLKFAEAPSLIIQMPRGKDFKLFKIFPSLELNITDLED 787
Qy 777 TPQCRICGLWAMECRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKPNVSLPKD 836
Db 788 TPQCRICGLWAMECRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKPNVSLPKD 847
Qy 837 LPDWMRHGCIQCNMELFAVICIETSHYVAFVKYKQDSAMLFPDSMADRGQNGFNI 896
Db 848 LPDWMRHGCIQCNMELFAVICIETSHYVAFVKYKQDSAMLFPDSMADRGQNGFNI 907
Qy 897 PQVTPCEVGEYLVKMSLEDLHSLDSRIQGCARRLLCDAYMCYQSPMTSLYK 949
Db 908 PQVTPCEVGEYLVKMSLEDLHSLDSRIQGCARRLLCDAYMCYQSPMTSLYK 960

RESULT 4

ADQ95916

ID ADQ95916 standard; protein; 953 AA.

XX AC ADQ95916;

XX DT 07-OCT-2004 (first entry)

XX DE T cell activation associated protein #47.

XX KW antiallergic; antiarthritic; antiaesthatic; antidiabetic; anti-HIV;
XX KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
KW gene therapy; T cell activation; diagnosis; autoimmune disease;
KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
KW bone-marrow transplant.

XX OS Homo sapiens.

XX PN WO2004058805-A2.

XX PD 15-JUL-2004.

XX PF 25-DEC-2003; 2003WO-JP016715.

XX

PR 26-DEC-2002; 2002JP-00376365.
PR 27-DEC-2002; 2002US-0436473P.
PR 28-APR-2003; 2003JP-00122113.
PR 28-APR-2003; 2003US-0465792P.
PR 21-OCT-2003; 2003JP-00360559.
PR 22-OCT-2003; 2003US-0512846P.

PA (ASAH-) ASahi Kasei Pharma Corp.

PI Matsuda A, Yoneta S;

XX WPI: 2004-593134/57.

DR N-PSDB; ADQ95915.

XX

PT New purified protein involved in T cell activation, useful for
PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
PT and infectious diseases.

PS Claim 1; SEQ ID NO 94; 2828pp; English.

XX The invention relates to purified proteins and genes encoding them, that
XX are involved in T cell activation (I) and has an amino acid deletion
CC substitution or addition in the amino acid sequences. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
CC marrow transplant. This sequence corresponds to a protein involved in T
XX cell activation.

SQ Sequence 953 AA;

Query Match 98.9%; Score 4980; DB 8; Length 953;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 947; Conservative 1; Mismatches 1; Indels 4; Gaps 4;

Qy 1 MSSGLWSQEKVTSFYWEERIFYLLOECSTVDKQTKLLKVPKSGIQYIQDRSVGHSRI 60

Db 1 MSSGLWSQEKVTSFYWEERIFYLLOECSTVDKQTKLLKVPKSGIQYIQDRSVGHSRI 60

Qy 61 PSAKKNQIQLKLEQPHAVLFVDZ-DVVEINEKTELLAINTCERESLFPKRNRLS 119

Db 61 PSAKKNQIQLKLEQPHAVLFVDZ-DVVEINEKTELLAINTCERESLFPKRNRLS 120

Qy 120 KGLQIDVGCVPVKQLESKEKPGVVRFRGPLLAERTVSGIFGVLELLEGRGOGFTDGV 179

Db 121 KGLQIDVGCVPVKQLESKEKPGVVRFRGPLLAERTVSGIFGVLELLEGRGOGFTDGV 180

Qy 180 YQKQLFQDCDECG-FVALDKLELIEDDDTALSDYAGPDTMQVELPPLLEINSRSLKG 238

Db 181 YQKQLFQDCDECGFVALDKLELIEDDDTALSDYAGPDTMQVELPPLLEINSRSLKV 240

Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDRFDGVLQCSFACVESTILLHN 297

Db 241 GETIESGTIVFCVLPKESLGYFVGVDMDNPIGNWDRFDGVLQCSFACVESTILLHN 300

Qy 298 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNR-RSELYFTLLNGSSV 356

Db 301 DIIPESVTQERRPPKLAFLMRGVDGKSSSHNKPATGSTSDPGNRSELFTLLNGSSV 360

Qy 357 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTFNRFHSLPFSL 416

Db 361 DSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTFNRFHSLPFSL 420

Qy 417 TKMPNTNGSIHSPUSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 476

Db 421 TKMPNTNGSIHSPUSLSAQSVMEELNTPAQESPLAMPNGSHGLEVGSIAEVKENPP 480

Qy 477 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKKAFLVKLSCRPDS 536

Db 481 FYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCAKKAFLVKLSCRPDS 540

QY 537 REASLOPVSNQIERCNSLAFGGVLSRVVETPPPKMEKSGLEIMIGKKGIQGHVNSCYL 596
Db 541 RFASLOPVSNQIERCNSLAFGGVLSRVVETPPPKMEKSGLEIMIGKKGIQGHVNSCYL 600
QY 597 DSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIMKLR 656
Db 601 DSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIMKLR 660
QY 657 KILEKVEAASGFTSEBKDPPEEFNLIFHILRLRVEPLLKIRSAGQKQVDCYFYQIFMEKNE 716
Db 661 KILEKVEAASGFTSEBKDPPEEFNLIFHILRLRVEPLLKIRSAGQKQVDCYFYQIFMEKNE 720
QY 717 KVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNITDLED 776
Db 721 KVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNITDLED 780
QY 777 TPRQCRIICGLAMYECCRECYDDPDISAGIKIQFCKTCNTQVHLHPKRLNHNKYNPVSILPKD 836
Db 781 TPRQCRIICGLAMYECCRECYDDPDISAGIKIQFCKTCNTQVHLHPKRLNHNKYNPVSILPKD 840
QY 837 LPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNGFNI 896
Db 841 LPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNGFNI 900
QY 897 PQTPTCPPEVGEYLKMSLEDLHLSDSRRIOGCARRLLCDAYMCMYQSPPTMSLYK 949
Db 901 PQTPTCPPEVGEYLKMSLEDLHLSDSRRIOGCARRLLCDAYMCMYQSPPTMSLYK 953

RESULT 5

ABB82783 ID ABB82783 standard; protein; 956 AA.

XX AC ABB82783;

XX DT 18-MAR-2003 (first entry)

XX DE Human CYLD polypeptide.

XX KW CD40; NF-kappaB essential modulator; NEMO; CYLD; immunosuppressive;
XX KW antiinflammatory; cytostatic; gene therapy; human.

XX OS Homo sapiens.

XX PN WO200292761-A2.

XX PD 21-NOV-2002.

XX PF 07-MAY-2002; 2002WO-US014570.

XX PR 08-MAY-2001; 2001US-00851673.

XX PA (IMMV) IMMUNEX CORP.

XX PI Derry JMJ, Fanslow WC, Dougall WC;

XX DR WPI; 2003-120669/11.

XX DR N-PSDB; ABV75394.

XX PT Identifying compounds that alter one or more biological activities of
XX PT CD40 by modulating the binding of NEMO and CYLD, useful for treating
XX PT disorders of the immune system, and inflammatory and cancer diseases.

XX PS Claim 1; Page 38-44; 48pp; English.

XX CC The invention relates to methods of identifying compounds that alter one/
XX CC more biological activities of CD40. One method involves screening for
XX CC molecules that modulate the binding of NF-kappaB essential modulator
XX CC (NEMO) and CYLD. The methods and compositions of the invention of
XX CC determining compounds that agonize or antagonize a CD40 signaling
XX CC activity, are useful for the further definition of CD40-mediated
XX CC signaling pathways, and for manipulation of CD40-mediated cellular
XX CC responses. They also provide therapeutic agents for treating disorders of

CC the immune system, and inflammatory and cancer diseases. The present
CC sequence represents the human CYLD polypeptide

XX SQ Sequence 956 AA;

Query Match 98.8%; Score 4971.5; DB 6; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLMSQEKVTSFYWEERIFYLLQECVSTDKQTKLLKVPKSGISQYIQRDSVGHRSI 60

Db 1 MSSGLMSQEKVTSFYWEERIFYLLQECVSTDKQTKLLKVPKSGISQYIQRDSVGHRSI 60

QY 61 PSAGKKNQIKLILQPHAVLVDE--DVVEINEKFTTELLAINTCBERFSLFNRRNLS 119

Db 61 PSAGKKNQIKLILQPHAVLVDEKDVVEINEKFTTELLAINTCBERFSLFNRRNLS 120

QY 120 KGIQIDVGCVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVVLEELGEGGQFTDGV 179

Db 121 KGIQIDVGCVKVQLRSGBEKFPGVVRFRGPLLAERTVSGIFFGVVLEELGEGGQFTDGV 180

QY 180 YQKQLFQCDDECG-FVALDKLEIEDDDTALESDDYAGPGDTMQVELPPLPINSRVSLKG 238

Db 181 YQKQLFQCDDECGFVALDKLEIEDDDTALESDDYAGPGDTMQVELPPLPINSRVSLKV 240

QY 239 GETIESGTIVFCVLPKESLGYFVGVDMPNPIGNWDGRFDGV-LCSFACVESTILLHIN 297

Db 241 GETIESGTIVFCVLPKESLGYFVGVDMPNPIGNWDGRFDGVLCSPFACVESTILLHIN 300

QY 298 DIIP---ESVTOERRPPKLAFMGRGVCDKSSSHNKPKATGSTSDPGRN-RSELYFTLNG 353

Db 301 DIIPALSESIVTOERRPPKLAFMGRGVCDKSSSHNKPKATGSTSDPGRNRSSELYFTLNG 360

QY 354 SSVDSQPQSKSKNTWYDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTERNFHSLP 413

Db 361 SSVDSQPQSKSKNTWYDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTTERNFHSLP 420

QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQESPPPLAMPFGNSHGLEVGSIAEYKE 473

Db 421 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQESPPPLAMPFGNSHGLEVGSIAEYKE 480

QY 474 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTDTGFRGTRYFTCALKKALFVKLKSQR 533

Db 481 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTDTGFRGTRYFTCALKKALFVKLKSQR 540

QY 534 PDSRFASLOPVSNQIERCNSLAFGGVLSRVVETPPPKMEKSGLEIMIGKKGIQGHVNS 593

Db 541 PDSRFASLOPVSNQIERCNSLAFGGVLSRVVETPPPKMEKSGLEIMIGKKGIQGHVNS 600

QY 594 CYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIM 653

Db 601 CYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTEIVNPLRIYGVYCATKIM 660

QY 654 KLRKILEKVEAASGFTSEBKDPPEEFNLIFHILRLRVEPLLKIRSAGQKQVDCYFYQIFME 713

Db 661 KLRKILEKVEAASGFTSEBKDPPEEFNLIFHILRLRVEPLLKIRSAGQKQVDCYFYQIFME 720

QY 714 KNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNITDL 773

Db 721 KNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSLIIQMPREGKDFKPKIFPSLELNITDL 780

QY 774 LEDTPRQCRIICGLAMYECCRECYDDPDISAGIKIQFCKTCNTQVHLHPKRLNHNKYNPVS 833

Db 781 LEDTPRQCRIICGLAMYECCRECYDDPDISAGIKIQFCKTCNTQVHLHPKRLNHNKYNPVS 840

QY 834 PKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNG 893

Db 841 PKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDSDMADRDGGQNG 900

QY 894 FNIPTPTCPPEVGEYLKMSLEDLHLSDSRRIOGCARRLLCDAYMCMYQSPPTMSLYK 949

Db 901 FNIPTPTCPPEVGEYLKMSLEDLHLSDSRRIOGCARRLLCDAYMCMYQSPPTMSLYK 956

RESULT 6

ADRI14489
 ID ADRI14489 standard; protein; 956 AA.
 AC ADRI14489;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human NF-kappaB pathway-associated protein SeqID490.
 XX
 KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
 KW antiarthritic; antirheumatic; gastrointestinal-gen; antiasthmatic;
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnary; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004065577-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004WO-US000798.
 XX
 PF 14-JAN-2003; 2003US-0440068P.
 PR 12-MAY-2003; 2003US-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Nadler SG, Neubauer MG, Feder JN, Carman J;
 XX
 DR WPI; 2004-562168/54.
 DR N-PSDB; ADRI14488.
 XX
 XX New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 PS Claim 6; SEQ ID NO 490; 237pp; English.
 XX
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
 CC gastrointestinal-gen, antiasthmatic, antiarteriosclerotic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnary activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, an inflammatory disorder
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,

CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction, infected
 CC proliferating disorders, cancers and HIV propagation in cells which
 CC is subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.
 XX
 SQ Sequence 956 AA;
 Query Match 98.8%; Score 4971.5; DB 8; Length 956;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;
 QY 1 MSSGLWSQKVTSPYWEERIFVLLQLQECSTVDKOTKLLKVPKSGISQYIQDRSVGHSRI 60
 DB 1 MSSGLWSQKVTSPYWEERIFVLLQLQECSTVDKOTKLLKVPKSGISQYIQDRSVGHSRI 60
 QY 61 PSAKGKKNQIGLIKLEPHAVLFDVE-DVVEINEKFTTELLAITNCEERSLFAKNRRLS 119
 DB 61 PSAKGKKNQIGLIKLEPHAVLFDVEKDVVEINEKFTTELLAITNCEERSLFAKNRRLS 120
 QY 120 KGLQIDVGCPCVKQLRSGEKPGVVRFRGPLAAERTVSGIFFGVELLEBGRGQGFDTGV 179
 DB 121 KGLQIDVGCPCVKQLRSGEKPGVVRFRGPLAAERTVSGIFFGVELLEBGRGQGFDTGV 180
 QY 180 YQKQLFQCDDECG-FVALDKLELEIEDDDTALESYAGPGDTMQVELPPELINSRVLKG 238
 DB 181 YQKQLFQCDDECGVFVALDKLELEIEDDDTALESYAGPGDTMQVELPPELINSRVLKV 240
 QY 239 GETIESGTIVFCDLPGKSLGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHIN 297
 DB 241 GETIESGTIVFCDLPGKSLGYFVGVDMDNPIGNWDGRPDGVQVLCSTFACVESTILLHIN 300
 QY 298 DIIP---ESVTOERRPPKLAFAKSRGVDKSGSSSHNKPATGSTDGNNR-RSELYFTLNG 353
 DB 301 DIIPALSESVOERRPPKLAFAKSRGVDKSGSSSHNKPATGSTDGNNRSELYFTLNG 360
 QY 354 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTENRHSPL 413
 DB 361 SSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTENRHSPL 420
 QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQSSPPLAMPNGSHGLEVSLAEVKE 473
 DB 421 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTPVQSSPPLAMPNGSHGLEVSLAEVKE 480
 QY 474 NPPFYGIWIRGIPGPGINEVLAGELEDEBCAGCTDGTFRGTRYFTCALKKALFVKLKSCR 533
 DB 481 NPPFYGIWIRGIPGPGINEVLAGELEDEBCAGCTDGTFRGTRYFTCALKKALFVKLKSCR 540
 QY 534 PDSRFASLOPVSNQIERCNSLAFGGYLSVEVENTPPKMEKEGLIMIGKKGIQGHVNS 593
 DB 541 PDSRFASLOPVSNQIERCNSLAFGGYLSVEVENTPPKMEKEGLIMIGKKGIQGHVNS 600
 QY 594 CYLDSTLFCFLAFSSVLDTVLLRPKKNDEYVYSETQELLRTIYNPLRIYGVYCATKIM 653
 DB 601 CYLDSTLFCFLAFSSVLDTVLLRPKKNDEYVYSETQELLRTIYNPLRIYGVYCATKIM 660
 QY 654 KLRKILEKVEAASGFTSEKDPPEEFNLPHHILRVESPLLKIRSAGQKQVDCFYQIFME 713
 DB 661 KLRKILEKVEAASGFTSEKDPPEEFNLPHHILRVESPLLKIRSAGQKQVDCFYQIFME 720
 QY 714 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLLIQMPFGCKDFKLFKKIPPSLELNTDL 773
 DB 721 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLLIQMPFGCKDFKLFKKIPPSLELNTDL 780
 QY 774 LEDTPRQCRICGGLAMVECRECYDDPDISAGIKQFCKTCTNTQVHLHPKRLNHNKYNPVS 833
 DB 781 LEDTPRQCRICGGLAMVECRECYDDPDISAGIKQFCKTCTNTQVHLHPKRLNHNKYNPVS 840
 QY 834 PKDLPDWDRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLFFDSMADRDGGONG 893
 DB 841 PKDLPDWDRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLFFDSMADRDGGONG 900

QY 894 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGARRLLCDAYMCMYQSPMTSLYK 949
DB 901 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGARRLLCDAYMCMYQSPMTSLYK 956

RESULT 7
ADQ95920
ID ADQ95920 standard; protein; 956 AA.
AC ADQ95920;
DT 07-OCT-2004 (first entry)
DE T cell activation associated protein #49.
XX
KW antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
KW gene therapy; T cell activation; diagnosis; autoimmune disease;
KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
KW bone-marrow transplant.
XX
OS Homo sapiens.
XX
XX
PN WO2004058805-A2.
XX
PD 15-JUL-2004.
XX
PF 25-DEC-2003; 2003WO-JP016715.
XX
PR 26-DEC-2002; 2002JP-00376365.
PR 27-DEC-2002; 2002US-0436473P.
PR 25-APR-2003; 2003JP-00122113.
PR 28-APR-2003; 2003US-0465792P.
PR 21-OCT-2003; 2003JP-00360559.
PR 22-OCT-2003; 2003US-0512846P.
XX
PA (ASAH-) ASahi Kasei PHARMA CORP.
PI Matsuda A, Yoneta S;
XX WPI; 2004-593134/57.
DR N-PSDB; ADQ95919.
XX
PT New purified protein involved in T cell activation, useful for
PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
PT and infectious diseases.
XX
PS Claim 1; SEQ ID NO 98; 2828pp; English.
XX
CC The invention relates to purified proteins and genes encoding them, that
CC are involved in T cell activation (I) and has an amino acid deletion,
CC substitution or addition in the amino acid sequences. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
CC marrow transplant. This sequence corresponds to a protein involved in T
CC cell activation.
XX
SQ Sequence 956 AA;

Query Match 98.7%; Score 4968.5; DB 8; Length 956;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 947; Conservative 1; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQKVTSPYWEERIFVLLQECSTDKQTKLLKVPKSGTQYIQDRSVGHSRI 60
DB 1 MSSGLWSQKVTSPYWEERIFVLLQECSTDKQTKLLKVPKSGTQYIQDRSVGHSRI 60
QY 61 PSAKGKKNQIGLKILPQPHAVLFVDE-DVVEINEKFTELLALTNCERFSLFKRNRLS 119

DB 61 PSAKGKKNQIGLKILPQPHAVLFVDEKDVVEINEKFTELLALTNCERFSLFKRNRLS 120
QY 120 KGLQIDVGCPCVKQLRSGEKFPVVRFRGPLAAERTVSGIFFGVVLEEGRGQGTGCV 179
DB 121 KGLQIDVGCPCVKQLRSGEKFPVVRFRGPLAAERTVSGIFFGVVLEEGRGQGTGCV 180
QY 180 YQKQLFOCEDCG-FVALDKLELIEDDDTALSDYAGDGTQVLPPLPINSRVSILK 238
DB 181 YQKQLFOCEDCGFVALDKLELIEDDDTALSDYAGDGTQVLPPLPINSRVSILK 240
QY 239 GETIESGTIVFCDLPGKESLGYFVGVDMDNPNIGMDGRFDGV-LCSFACVESTILLHN 297
DB 241 GETIESGTIVFCDLPGKESLGYFVGVDMDNPNIGMDGRFDGV-LCSFACVESTILLHN 300
QY 298 DIIP---ESVTQERRPPKLAFMGRGVGDKGSSSHNPKATGSTSDPQNR-RSELYFTLNG 353
DB 301 DIIPALSESVTQERRPPKLAFMGRGVGDKGSSSHNPKATGSTSDPQNRSELFTLNG 360
QY 354 SSVDSQFQSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPNVSLTTERFHSPL 413
DB 361 SSVDSQFQSKSNWTYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPNVSLTTERFHSPL 420
QY 414 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTAPVQESPLAMPNGSHGLEVSLAEVKE 473
DB 421 FSLTKMPTNGSIGHSPLSLSAQSVMEELNTAPVQESPLAMPNGSHGLEVSLAEVKE 480
QY 474 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTFRTYFTCALKKALFVKLXSCR 533
DB 481 NPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTFRTYFTCALKKALFVKLXSCR 540
QY 534 PDSRFASLPQVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKKGIGHYNS 593
DB 541 PDSRFASLPQVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKKGIGHYNS 600
QY 594 CYLDSTLFCLEAFSSVLDTVLLRPKKNVVEYSETQELLRTETVNPRLIYGVYCATKIM 653
DB 601 CYLDSTLFCLEAFSSVLDTVLLRPKKNVVEYSETQELLRTETVNPRLIYGVYCATKIM 660
QY 654 KLRKILEKVEAASGFTSEKDPPEFLNLPFHILRVEPLLKIRSAQKQVQDCYFYQIFME 713
DB 661 KLRKILEKVEAASGFTSEKDPPEFLNLPFHILRVEPLLKIRSAQKQVQDCYFYQIFME 720
QY 714 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLIIQMPRGKDFKLFKKIPPSLELNTDL 773
DB 721 KNEKVGPTIQQLLEWSFINSNLKFAEAPSCLIIQMPRGKDFKLFKKIPPSLELNTDL 780
QY 774 LEDTPRQCRICGGLAMYECRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVS 833
DB 781 LEDTPRQCRICGGLAMYECRECYDDPDISAGKIKQFCKTCNTQVHLHPKRLNHNKYNPVS 840
QY 834 PKDLPDMDWRHGCTPCQNMELFVLCITSHYVAFVKYKDDSAWLFDDSMADRDGGONG 893
DB 841 PKDLPDMDWRHGCTPCQNMELFVLCITSHYVAFVKYKDDSAWLFDDSMADRDGGONG 900
QY 894 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGARRLLCDAYMCMYQSPMTSLYK 949
DB 901 FNIPQVTPCEVGEYLKMSLEDLHSLDSRRIOGARRLLCDAYMCMYQSPMTSLYK 956

RESULT 8
AAU23747
ID AAU23747 standard; protein; 739 AA.
XX
AC AAU23747;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human enzyme polypeptide #833.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;

KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
XX nephrotropic; anticoagulant.

OS Homo sapiens.

XX WO200155301-A2.

PN

XX PD

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001239.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190078P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226682P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.

PR 01-DEC-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.

PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.

PR	11-DEC-2000; 2000US-0254097P.	QY	661	KVEAASGFTSEKDPPEFLNLFHHILRVPELILKIRSAGQVQDCYFYQIFMEKNEKVG	720
PR	05-JAN-2001; 2001US-0259678P.	Db	451	KVEAASGFTSEKDPPEFLNLFHHILRVPELILKIRSAGQVQDCYFYQIFMEKNEKVG	510
PA	(HUMA-) HUMAN GENOME SCI INC.	QY	721	PTIQQLLEWSFINSNLKFAEAPSCLIIQMPFRFGDKFLFKKIFPSLELNTDLEDTPRQ	780
XX	Rosen CA, Barash SC, Ruben SM;	Db	511	PTIQQLLEWSFINSNLKFAEAPSCLIIQMPFRFGDKFLFKKIFPSLELNTDLEDTPRQ	570
XX	WPI, 2001-465566/50.	QY	781	CRICGGLAMVECRECYDDPDISAGIKQFCCTNTQVHLHPKRLNHKYNPVSPLPKDLPDW	840
DR	N-PSDB; AAS41617.	Db	571	CRICGGLAMVECRECYDDPDISAGIKQFCCTNTQVHLHPKRLNHKYNPVSPLPKDLPDW	630
XX	Novel polypeptides and polynucleotides useful for diagnosing, preventing,	QY	841	DWRHGICPCQNMELFAVLCTIETSHYVAFVKYKDDSNLFPDSMADRDGGONGNIPOVT	900
PT	treating neural, immune system, muscular, reproductive, pulmonary,	Db	631	DWRHGICPCQNMELFAVLCTIETSHYVAFVKYKDDSNLFPDSMADRDGGONGNIPOVT	690
PT	cardiovascular, renal, proliferative disorders and cancerous diseases.	QY	901	PCPEVGEYLKMSLEDLHSDRSRIQGCARRLLCDAYCMYQSPMTSLYK	949
XX	Claim 11; SEQ ID NO 1743; 1180pp; English.	Db	691	PCPEVGEYLKMSLEDLHSDRSRIQGCARRLLCDAYCMYQSPMTSLYK	739
CC	The present invention relates to the isolation of novel human enzyme	RESULT 9			
CC	polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences	AA919154			
CC	encoding them. The enzyme polypeptides of the invention may comprise the	ID	AA919154	standard; protein; 731 AA.	
CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,	XX	AA919154		
CC	isomerases or ligases. The sequences of the invention are useful in the	AC	AA919154		
CC	diagnosis, treatment, prevention and/or prognosis of a wide range of	XX	AA919154		
CC	disorders including hyperproliferative disorders (e.g. cancer),	DT	19-JUL-2000	(first entry) - late	
CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.	XX			
CC	arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic	DE		Human cytoskeleton associated protein 9 (CYSXP-9).	
CC	disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),	XX			
CC	cardiovascular disorders (e.g. atherosclerosis), blood-related disorders	KW		Cytoskeleton associated protein; CYSXP-9; cancer; proliferative;	
CC	(e.g. haemophilia), reproductive disorders (e.g. infertility) and	KW		autoimmunity; inflammatory, vesicle trafficking; neurological;	
CC	infectious disorders (e.g. influenza). The polynucleotides of the	KW		cardiovascular; cell motility; reproductive; muscle disorder.	
CC	invention can also be used in gene therapy. AAU22915-AAU23814 represent	XX			
CC	the novel human enzyme polypeptides of the invention. Note: The sequence	OS		Homo sapiens.	
CC	data for this patent did not form part of the printed specification, but	XX			
CC	was obtained in electronic format directly from WIPO at	PH			
CC	ftp.wipo.int/pub/published_pct_sequences	FT			
XX	Sequence 739 AA;	FT			
SQ		FT			
QY	184 QLFQCDDECG-FVALDKLELIEDDDTALESYAGPGDTMQVELPPLNRSVLSKGETI	FT			
Db	1 QLFQCDDECG-FVALDKLELIEDDDTALESYAGPGDTMQVELPPLNRSVLSK-----	FT			
QY	243 ESGTVIFCDVLPKESLGIFVGVDMNDPIGNWDRPDGV-LCSFACVESTILLHNDIIP	FT			
Db	56 -----DNPIGNWDRPDGV-LCSFACVESTILLHNDIIP	FT			
QY	302 ESVTQERRPPKLAFLMSRGVKGSSSHNKPATGTSDFGNR-RSELFYTLNGSSVDSQP	FT			
Db	91 ESVTQERRPPKLAFLMSRGVKGSSSHNKPATGTSDFGNRSELFTYTLNGSSVDSQP	FT			
QY	361 QSKSKNTWIDEVADPAKSLTEISTDFDRSPPLQPPVNSLTENRPHSLPFLTKMP	FT			
Db	151 QSKSKNTWIDEVADPAKSLTEISTDFDRSPPLQPPVNSLTENRPHSLPFLTKMP	FT			
QY	421 NTNGSIGHSPLSLSAQSVMEELNTPVQESPLAMPNGSHGLEVGSIAEVKENPPFYGV	FT			
Db	211 NTNGSIGHSPLSLSAQSVMEELNTPVQESPLAMPNGSHGLEVGSIAEVKENPPFYGV	FT			
QY	481 IRWIGQPPGLNEVLAGLEDEDCAGCTDGTFRGTRYFTCALKKALFVKLKSCRPPDSRFAS	FT			
Db	271 IRWIGQPPGLNEVLAGLEDEDCAGCTDGTFRGTRYFTCALKKALFVKLKSCRPPDSRFAS	FT			
QY	541 LQPVSNQIERCNSLAFGGVLSVWBEPTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSLT	FT			
Db	331 LQPVSNQIERCNSLAFGGVLSVWBEPTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSLT	FT			
QY	601 FCLFAPSSVLDTVLLRPKEKNDVEYVETQELLRTIENVPLRIYGVVCAATKMKLRKILE	FT			
Db	391 FCLFAPSSVLDTVLLRPKEKNDVEYVETQELLRTIENVPLRIYGVVCAATKMKLRKILE	FT			

PS Claim 8; SEQ ID NO 18843; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dr primer and an oligonucleotide complementary to the

CC complementary strand of a polynucleotide which comprises one of the 5602

CC nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893

CC represent human amino acid sequences; and AAH13629 to AAH13632 represent

CC oligonucleotides, all of which are used in the exemplification of the

CC present invention

XX Sequence 731 AA;

SEQ

Query Match 76.6%; Score 3855; DB 4; Length 731;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 727; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 221 MOVLPPLPPLINRSVLSKVGTTESGTVIFCDVLPKGESLGYFVGVDMDNPNIGNWGRFDG 280

DB 1 MOVLPPLPPLINRSVLSKVGTTESGTVIFCDVLPKGESLGYFVGVDMDNPNIGNWGRFDG 60

QY 281 V-LCSFACVESTILLHINDIIPESVTOERRPPKLAFMSRGVGDGSGSHNKPATGSTD 339

DB 61 VQLCSFACVESTILLHINDIIPESVTOERRPPKLAFMSRGVGDGSGSHNKPATGSTD 120

QY 340 PGNR-RSELFYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTETSTDFDRSSPPLQPP 398

DB 121 PGNRSELSELFYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTETSTDFDRSSPPLQPP 180

QY 399 PVNSLTENRPHSLPPLTKMPTNGSIGHSPLSLSAQSVMEELNAPVQESPLAMPFG 458

DB 181 PVNSLTENRPHSLPPLTKMPTNGSIGHSPLSLSAQSVMEELNAPVQESPLAMPFG 240

QY 459 NSHGLEVGSIAEVKENPPYGVIRWTGPPGLNEVLAGLEDEACAGCTDGTFRGTRYFT 518

DB 241 NSHGLEVGSIAEVKENPPYGVIRWTGPPGLNEVLAGLEDEACAGCTDGTFRGTRYFT 300

QY 519 CALKKALFVKLAKSRPDSRFASIQPVSNQIERCNSLAFGGLSEVVVEENTPPKWEKGLE 578

DB 301 CALKKALFVKLAKSRPDSRFASIQPVSNQIERCNSLAFGGLSEVVVEENTPPKWEKGLE 360

QY 579 IMIGKKGIQGHYNSCYLSDTLFCLFAPSSVLDTVLLRPKEKNDVYVSETQELLRTIV 638

DB 361 IMIGKKGIQGHYNSCYLSDTLFCLFAPSSVLDTVLLRPKEKNDVYVSETQELLRTIV 420

QY 639 NPLRIYGYCATKIMKRLKILEKVEAASGFTSEKDPBEEFLNLFHILRVEPLLKIRSA 698

DB 421 NPLRIYGYCATKIMKRLKILEKVEAASGFTSEKDPBEEFLNLFHILRVEPLLKIRSA 480

QY 699 GQKVQDCYQYQFMENKNEKVGPTIQOLLEWSPFNSLNKFAEAPSCLIIQMPRFQDFKL 758

DB 481 GQKVQDCYQYQFMENKNEKVGPTIQOLLEWSPFNSLNKFAEAPSCLIIQMPRFQDFKL 540

QY 759 FKKIPPSLENTDLDLTPROCRICGGIAMVRECYDDPDISAGKIKQFCKTCTQVH 818

DB 541 FKKIPPSLENTDLDLTPROCRICGGIAMVRECYDDPDISAGKIKQFCKTCTQVH 600

QY 819 LHPKRLNHNKYNPVSPLPKDLPDMDWRHGCCIPQNMELFAVLCTIETSHYVAFVKYGDSDSAW 878

DB 601 LHPKRLNHNKYNPVSPLPKDLPDMDWRHGCCIPQNMELFAVLCTIETSHYVAFVKYGDSDSAW 660

QY 879 LFPDSMADRDGGONGFNIPQVTPCPVEGYLKMSLDLHSLDRRIQGCARRLLCDAYMC 938

DB 661 LFPDSMADRDGGONGFNIPQVTPCPVEGYLKMSLDLHSLDRRIQGCARRLLCDAYMC 720

QY 939 MYQSPMTSLYK 949

DB 721 MYQSPMTSLYK 731

RESULT 11

AAH39254

ID AAH39254 standard; protein; 685 AA.

XX

AC AAH39254;

XX

DT 22-OCT-2001 (first entry) - *late*

XX

DE Human polypeptide SEQ ID NO 2399.

XX

KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

XX

OS Homo sapiens.

XX

XX WO200153312-A1.

XX

PD 26-JUL-2001.

XX

XX 26-DEC-2000; 2000WO-US034263.

XX

XX 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-0052317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX

XX (HYSE-) HYSEQ INC.

XX

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

PI Zhou P, Goodrich R, Drmanac RT;

XX

XX WPI; 2001-442253/47.

DR N-PSDB; AA158410.

XX

XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

XX

PS Example 4; SEQ ID NO 2399; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAH38642-AAH42213) with neotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and

CC	C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification	
XX		
SQ	Sequence 685 AA;	
	Query Match 72.2%; Score 3636; DB 4; Length 685;	
	Best Local Similarity 99.7%; Pred. No. 0;	
	Matches 683; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
Qy	267 MDNPIGNWGRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRGVDGKGS 325	
Db	1 MDNPIGNWGRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRGVDGKGS 60	
Qy	326 SSHNPKATGSTDGPNR-RSELFVTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEI 384	
Db	61 SSHNPKATGSTDGPNRSELFTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEI 120	
Qy	395 STDFDRSSPPLQPPVNSLTNNRPHSLPFLSKMPNTNGSIGHSPLSLSAQSVMEELNT 444	
Db	121 STDFDRSSPPLQPPVNSLTNNRPHSLPFLSKMPNTNGSIGHSPLSLSAQSVMEELNT 180	
Qy	445 APVQESPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPPGLEVLAGELEDECA 504	
Db	181 APVQESPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPPGLEVLAGELEDECA 240	
Qy	505 GCTDGTFRGTRYFTCALKKALFVKLSCRPDSRFASLQPVSNQIERCNSLAFGGYLSVV 564	
Db	241 GCTDGTFRGTRYFTCALKKALFVKLSCRPDSRFASLQPVSNQIERCNSLAFGGYLSVV 300	
Qy	565 BENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTLFCFLAFSSVLDVLLRPKEKNV 624	
Db	301 BENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTLFCFLAFSSVLDVLLRPKEKNV 360	
Qy	625 YYSETQELLRTIEVNLPLRYGVVCAIKMKLRKILEKVEAASGFTSEKDPPEEFNLPH 684	
Db	361 YYSETQELLRTIEVNLPLRYGVVCAIKMKLRKILEKVEAASGFTSEKDPPEEFNLPH 420	
Qy	685 HILRVEPLLKIRSAQKQVDCYFYQIFMEKNEKVGVPYTTIQLLEWSFINSNLKFAEAPSC 744	
Db	421 HILRVEPLLKIRSAQKQVDCYFYQIFMEKNEKVGVPYTTIQLLEWSFINSNLKFAEAPSC 480	
Qy	745 LIIQWPRFGKPKLFPKIPPSLELNITDLETPRQCRICGGLAMVECRECYDDPDISAG 804	
Db	481 LIIQWPRFGKPKLFPKIPPSLELNITDLETPRQCRICGGLAMVECRECYDDPDISAG 540	
Qy	805 KIKQFCKTCTQVHLHPKELNHYNPVSLPKDLPDMDWRHGCIPQNMELFAVLCTETSH 864	
Db	541 KIKQFCKTCTQVHLHPKELNHYNPVSLPKDLPDMDWRHGCIPQNMELFAVLCTETSH 600	
Qy	865 YVAFVKYKDDSAWLFFDSMDRGGGNGFNIPQVTPCPVEGVEYLKMSLEDLHSLDSRI 924	
Db	601 YVAFVKYKDDSAWLFFDSMDRGGGNGFNIPQVTPCPVEGVEYLKMSLEDLHSLDSRI 660	
Qy	925 QGCARLLCDAYMCMYQSTMSLYK 949	
Db	661 QGCARLLCDAYMCMYQSTMSLYK 685	
RESULT 12		
ID	AAM41040	
XX	AAM41040 standard; protein; 698 AA.	
XX		
AC	AAM41040;	
XX		
DT	22-OCT-2001 (first entry)	
XX	Human polypeptide SEQ ID NO 5971.	
DE		
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	

KW	leukaemia.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	26-DEC-2000; 2000WO-US034263.	
XX		
PR	23-DEC-1999; 99US-00471275.	
PR	21-JAN-2000; 2000US-00498725.	
PR	25-APR-2000; 2000US-0052317.	
PR	20-JUN-2000; 2000US-00598042.	
PR	19-JUL-2000; 2000US-00620312.	
PR	03-AUG-2000; 2000US-00653450.	
PR	14-SEP-2000; 2000US-00662191.	
PR	19-OCT-2000; 2000US-00693036.	
PR	29-NOV-2000; 2000US-00727344.	
XX	(HYSE-) HYSEQ INC.	
PA		
XX		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	
PI	Zhou P, Goodrich R, Drmanac RT;	
XX		
DR	WPI; 2001-442253/47.	
DR	N-PSDB; AAI60196.	
XX		
XX	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.	
PS	Example 2; SEQ ID NO 5971; 10078pp; English.	
XX		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification	
XX		
SQ	Sequence 698 AA;	
	Query Match 70.8%; Score 3563; DB 4; Length 698;	
	Best Local Similarity 97.8%; Pred. No. 5.2e-314;	
	Matches 680; Conservative 2; Mismatches 7; Indels 6; Gaps 6;	
Qy	261 YFVGVDMDNPIGNWGRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRG 319	
Db	4 YFVGVDMDNPIGNWGRFDGV-LCSFACVESTILLHINDIIPESVTQERRPPKLAFLMRG 63	
Qy	320 VDGKSSSHNPKATGSTDGPNR-RSELFVTLNGSSVDSQPSQSKNTWYIDEVAEDPA 378	
Db	64 VDGKSSSHNPKATGSTDGPNRSELFTLNGSSVDSQPSQSKNTWYIDEVAEDPA 123	
Qy	379 KSLTEISTDFDRSSPPLQPPVNSLTNNRPHSLPFLSKMPNTNGSIGHSPLSAQSV 438	
Db	124 KSLTEISTDFDRSSPPLQPPVNSLTNNRPHSLPFLSKMPNTNGSIGHSPLSAQSV 183	
Qy	439 MEELNTAPVQESPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPPGLEVLAGE 498	
Db	184 MEELNTAPVQESPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPPGLEVLAGE 243	
Qy	499 LEDECAGCTDGTGR-GTRYFTCALKKALFVKLSCRPDSRFASLQPVSNQIERCNSLA-F 556	

Db	244	LEDEAGCTDGTFRGTRYFTCALKKALFVKLKSRCDFSRFASIQPVNSQIERCNSLAIW	303		
Qy	557	GGYLSEVVEENTP-PKMEKEGLEIMIGKKGIQGHYNSCYLDSTLFCCLFAPSSVLDTVLL	615		
Db	304	EAYLSEVVEENTPTQWEKEGLEIMIGKKGIQGHYNSCYLDSTLFCCLFAPSSVLDTVLL	363		
Qy	616	RPKEKNDEYYSEYEQELLRTIIVNPLRIYGVVCAKIMKRLKILEKVEAASGFTSEKDP	675		
Db	364	RPKEKNDEYYSEYEQELLRTIIVNPLRIYGVVCAKIMKRLKILEKVEAASGFTSEKDP	423		
Qy	676	EEFLNLFFHLIRVEPLLKIRSAQKQVQDCYFYQIFMEKNEKVGVPITIQOLLEWSFNSN	735		
Db	424	EEFLNLFFHLIRVEPLLKIRSAQKQVQDCYFYQIFMEKNEKVGVPITIQOLLEWSFNSN	483		
Qy	736	LKFAEAPSCLIIQMPRFGKDFKFKIIPSPLELNITDLEDTPRQCRICGGLAWYECREC	795		
Db	484	LKFAEAPSCLIIQMPRFGKDFKFKIIPSPLELNITDLEDTPRQCRICGGLAWYECREC	543		
Qy	796	YDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVSLPKDLPDWDWRHGCIPQNNMELF	855		
Db	544	YDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVSLPKDLPDWDWRHGCIPQNNMELF	603		
Qy	856	AVLCIETSHVAFVKYKGDSSAWLFFDSMADRGQNGFNIPQVTPCPEVGEYILKMSLED	915		
Db	604	AVLCIETSHVAFVKYKGDSSAWLFFDSMADRGQNGFNIPQVTPCPEVGEYILKMSLED	663		
Qy	916	LHSLDSRRIQGCARRLLCDA-YMCWYQSPMTSLYK	949		
Db	664	LHSLDSRRIQGCARRLLCDAIYVPCQSPMTSLYK	698		
RESULT 13					
AAU23211					
ID	AAU23211 standard; protein; 558 AA.				
AC	AAU23211;				
XX	(first entry)				
DT	18-DEC-2001				
DE	Novel human enzyme polypeptide #297.				
XX	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;				
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;				
KW	autoimmune disorder; neurological disorder; metabolic disorder;				
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;				
KW	blood-related disorder; infectious disorder; cytostatic; anti arthritic;				
KW	nephrotropic; anticoagulant.				
OS	Homo sapiens.				
XX	WO200155301-A2.				
XX	02-AUG-2001.				
XX	17-JAN-2001; 2001WO-US001239.				
XX	31-JAN-2000; 2000US-0179065P.				
PR	04-FEB-2000; 2000US-0180628P.				
PR	24-FEB-2000; 2000US-0184664P.				
PR	02-MAR-2000; 2000US-0186350P.				
PR	16-MAR-2000; 2000US-0189874P.				
PR	17-MAR-2000; 2000US-0190076P.				
PR	18-APR-2000; 2000US-0198123P.				
PR	19-MAY-2000; 2000US-0205515P.				
PR	07-JUN-2000; 2000US-0209467P.				
PR	28-JUN-2000; 2000US-0214886P.				
PR	30-JUN-2000; 2000US-0215135P.				
PR	07-JUL-2000; 2000US-0216647P.				
PR	07-JUL-2000; 2000US-0216880P.				
PR	11-JUL-2000; 2000US-0217487P.				
PR	14-JUL-2000; 2000US-0217496P.				
PR	14-JUL-2000; 2000US-0218290P.				

PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUNA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465566/50.
DR N-PSDB; AAS41081.
XX
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
PS Claim 11; SEQ ID NO 1207; 1180pp; English.
XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders, including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and

CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 558 AA;
Query Match 54.7%; Score 2755; DB 4; Length 558;
Best Local Similarity 93.3%; Pred. No. 1.le-240;
Matches 539; Conservative 1; Mismatches 4; Indels 34; Gaps 5;
QY 1 MSSGLWSQKVTSPYWEERIFYLLQECVTDKQTKLLKVPKSGISQYITODRSVGHRI 60
DB 8 MSSGLWSQKVTSPYWEERIFYLLQECVTDKQTKLLKVPKSGISQYITODRSVGHXSI 67
QY 61 PSAKGGKQIGIKILEOPHAVLFVDE-DVVEINEKEFTELLAITNCEERSLFKNRRLS 119
DB 68 PSAKGGKQIGIKILEOPHAVLFVDEKVVINEKEFTELLAITNCEERSLFKNRRLS 127
QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFFGVLELLEGRGQGTGTV 179
DB 128 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFFGVLELLEGRGQGTGTV 187
QY 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPGDTMQVELPPLPPIEINRSVLKG 238
DB 188 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPGDTMQVELPPLPPIEINRSVLK- 246
QY 239 GETIESGTWIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 297
DB 247 -----DNPIGNWDGRFDGVQLCSFACVESTILLHIN 277
QY 298 DIIPESVTQERRPPKLAFMRSRGVGDKGSSHNPKATGSTSDPCGNR-RSELFYTLNGSSV 356
DB 278 DIIPESVTQERRPPKLAFMRSRGVGDKGSSHNPKATGSTSDPCGNRNRSELFYTLNGSSV 337
QY 357 DSQPSQSKNNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNRRFHSLPFSL 416
DB 338 DSQPSQSKNNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNRRFHSLPFSL 397
QY 417 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQSPPLAMPNGSHGLEVGSIAEVKENPP 476
DB 398 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQSPPLAMPNGSHGLEVGSIAEVKENPP 457
QY 477 FYGVIRWIGOPGLNEVLGLELEDECACTDGTFRGTRTYFTCAKKALFVKLKSCRPS 536
DB 458 FYGVIRWIGOPGLNEVLGLELEDECACTDGTFRGTRTYFTCAKKALFVKLKSCRPS 517
QY 537 RFASLPVSNQIERCNSLAFGGYLSEVVEENTPPKMEK 574
DB 518 RFASLPVSNQIERCNSLAFGGYLSEVVEENTPPSSE 555
RESULT 14
AAB95719
ID AAB95719 standard; protein; 476 AA.
XX
AC AAB95719;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:18587.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.

```

XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 19587; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dt primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 476 AA;
XX
XX Query Match 46.5%; Score 2343; DB 4; Length 476;
XX Best Local Similarity 98.1%; Pred. No. 2.3e-203;
XX Matches 454; Conservative 2; Mismatches 3; Indels 4; Gaps 3;
XX
QY 221 MQVELPPLEINRSVSLKGGTIESGTVIFCDVLPKGESLGYPGVDMNDPNIQNWDRFGD 280
DB 1 MQVELPPLEINRSVSLKGGTIESGTVIFCDVLPKGESLGYPGVDMNDPNIQNWDRFGD 60
QY 281 V-LCSFACVESTILLHINDIIPESVTQERRPKLAFMSRGGVKGSSSHNKPATGSTSD 339
DB 61 VOLCSFACVESTILLHINDIIPESVTQERRPKLAFMSRGGVKGSSSHNKPATGSTSD 120
QY 340 PGNR-RSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPP 398
DB 121 PGNRNSRSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPP 180
QY 399 PVSNTLTENRPSLTPSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPLAMP 458
DB 181 PVSNTLTENRPSLTPSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPLAMP 240
QY 459 NSHGLEVGSLAEVKENPPFYGVIRWTGPPPLNEVLGLAELEDECACTDGTGFRGTRYFT 518
DB 241 NSHGLEVGSLAEVKENPPFYGVIRWTGPPPLNEVLGLAELEDECACTDGTGFRGTRYFT 300
QY 519 CALKALFKVLKSCRDPDSFASLQPVSNQIERNCSLAFGGYLSVVVEENTPPKMEKGLE 578
DB 301 CALKALFKVLKSCRDPDSFASLQPVSNQIERNCSLAFGGYLSVVVEENTPPKMEKGLE 360

```

```

QY 579 IMIGKKKGIQGHVNSCYLDSTLFCFLPAPSSVLDTVLLRPKEKNDVEYYSQELLRTETV 638
DB 361 IMIGKKKGIQGHVNSCYLDSTLFCFLPAPSSVLDTVLLRPKEKNDVEYYSQELLRTETV 420
QY 639 NPLRIYGVVCAATKIMLKRLKILEKVEAASGGTSEKDPPEFLNI 681
DB 421 NPLRIYGVVCAATKIMLKRLKILEKVEAASGGTSEKDPPEFLNI 461

RESULT 15
ABB89233
ID ABB89233 standard; protein; 261 AA.
XX
XX ABB89233;
XX AC
XX DT 24-MAY-2002 (first entry)
XX
XX Human polypeptide SEQ ID NO 1609.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX
XX Homo sapiens.
XX
XX WO200190304-A2.
XX
XX 29-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-US016450.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-122018/16.
XX
XX N-PSDB; ABL89642.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX prevention of neural, immune system, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders.
XX
XX Claim 11; SEQ ID NO 1609; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (ABB9040-ABB90444) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 261 AA;
XX
XX Query Match 25.3%; Score 1275; DB 5; Length 261;
XX Best Local Similarity 98.7%; Pred. No. 1.1e-106;
XX Matches 234; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy	653	MKLRKILEKVEAASGFTSEKDPPEFLNLFHHLRVEPLLKIRSAGQKQVDCYFYQIEM	712
Db	1	1 MKLRKILEKVEAASGFTSEKDPPEFLNLFHHLRVEPLLKIRSAGQKQVDCYFYQIEM	60
Qy	713	EXNEKVGVPITIOQLLEWSFINSNLKEAEAPSCLIIOMPRGKDFKLFKKIFPSLELNITD	772
Db	61	EXNEKVGVPITIOQLLEWSFINSNLKEAEAPSCLIIOMPRGKDFKLFKKIFPSLELNITD	120
Qy	773	LLEDTPROCRICGGLAMYECRECYDDPDISAGKIQFCKTCNTQVHLHPKRLNHNKYNPVS	832
Db	121	LLEDTPROCRICGGLAMYECRECYDDPDISAGKIQFCKTCNTQVHLHPKRLNHNKYNPVS	180
Qy	833	LPKDLPDWDRHGCIPQNMELFAVLCTIETSHYVAFVYKGDSDAWLFFDSMADRDG	889
Db	181	LPKDLPDWDRHGCIPQNMELFAVLCTIETSHYVAFVYKGDSDAWLFFDSMADRDG	237

Search completed: April 18, 2005, 15:23:39
 Job time : 79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:32:50 ; Search time 6177 Seconds
(without alignments)
5847.983 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCYQSPYMSLYK 949

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool_p/US09671687/runat.18042005.115018.17139/app_query.fasta_1.1095
-DB=EST-QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bts -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09671687 @CGN 1 5180 @runat.18042005.115018.17139 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc2:*
4: gb_hc3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4494	89.3	3496	3 AF161542	Homo sapi
2	4313	85.7	2862	9 AY406374	Homo sapi
3	4102	81.5	2730	9 AY406376	Mus muscu
4	3768	74.9	2862	9 AY406375	Pan trogl
5	2952	58.6	2220	3 AK039054	Mus muscu
6	2769	55.0	3137	3 BC028885	Mus muscu
7	1572.5	31.2	1063	5 BQ433523	AGENCOURT
8	1570	31.2	1039	4 BM457960	AGENCOURT
9	1459.5	29.0	1478	3 AK042764	Mus muscu

10	1435	28.5	938	4	BI666276	603290787
11	1426	28.3	2218	3	AK048183	Mus muscu
12	1416.5	28.1	1133	4	BM480127	AGENCOURT
13	1384	27.5	812	6	CA512526	UI-R-FJO-
14	1363	27.1	1892	3	BC024596	Mus muscu
15	1340	26.6	837	4	BG867631	602787414
16	1302	25.9	816	5	BQ442079	UI-N-EXO-
17	1288	25.6	721	5	BP162259	BP162259
18	1286.5	25.6	889	5	BUS05487	AGENCOURT
19	1280	25.4	783	4	BM016881	603643383
20	1268.5	25.2	715	5	CN157036	944487 MA
21	1267.5	25.2	840	5	BP144628	BP144628
22	1262	25.1	837	1	AU137267	AU137267
23	1235	24.5	726	7	CR772310	KXP2p468L
24	1233	24.5	717	7	CF750356	UI-M-HJO-
25	1215	24.1	758	1	AU122742	AU122742
26	1202.5	23.9	791	7	CN461343	UI-M-HNO-
27	1198	23.8	743	6	CD351332	UI-M-GIO-
28	1170	23.2	677	6	CD626856	56076837J
29	1165	23.1	648	1	AL707441	KXP2p686M
30	1165	23.1	697	4	BM724143	UI-E-EOI-
31	1160.5	23.1	781	7	CK636353	UI-M-HNO-
32	1158	23.0	826	7	CK597829	AGENCOURT
33	1152	22.9	708	6	CB466941	732655 MA
34	1152	22.9	794	2	BF141863	601791346
35	1151	22.9	770	4	BI906770	603064511
36	1139.5	22.6	785	6	CB235676	AGENCOURT
37	1128	22.4	906	5	BU461394	603774393
38	1127.5	22.4	907	5	BQ438227	AGENCOURT
39	1118	22.2	702	2	BB648023	BB648023
40	1117.5	22.2	755	7	CO404877	AGENCOURT
41	1107.5	22.0	723	6	CD559213	AGENCOURT
42	1105.5	22.0	1134	5	BM904688	AGENCOURT
43	1102	21.9	637	2	AW985722	uf85b05.Y
44	1094	21.7	639	7	CN704295	E0485H10-
45	1092.5	21.7	1365	3	AY383658	Rattus no

ALIGNMENTS

RESULT 1	AF161542	Homo sapiens HSPC057 mRNA, complete cds.	linear	HTC 22-MAY-2001
LOCUS	AF161542	3496 bp	mrna	
DEFINITION	Homo sapiens HSPC057 mRNA, complete cds.			
ACCESSION	AF161542			
VERSION	AF161542.1	GI:6841351		
KEYWORDS	HTC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 3496)			
AUTHORS	Zhang, Q.H., Ye, M., Wu, X.Y., Ren, S.X., Zhao, M., Zhao, C.J., Fu, G., Shen, Y., Fan, H.Y., Lu, G., Zhong, M., Xu, X.R., Han, Z.G., Zhang, J.M., Tao, J., Huang, Q.H., Zhou, J., Hu, G.X., Gu, J., Chen, S.J. and Chen, Z.			
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells			
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)			
MEDLINE	20499367			
PUBMED	11042152			
REFERENCE	2 (bases 1 to 3496)			
AUTHORS	Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.			
TITLE	Human full length cDNA cloned from cd34+ stem cells			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 3496)			
AUTHORS	Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-MAY-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China			

FEATURES

Location/Qualifiers
 1..3496 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CBFA0B03"
 /tissue_type="blood"
 1517..3064
 CDS
 /codon_start=1
 /product="HSPC057"
 /protein_id="AAF29029.1"
 /db_xref="GI:6841352"
 /translation="MENTNGSIHSPLSAOSVMEELNTAPVQBSPPPLAMPNGSHG
 LEVGSIAEVRKPNPPFGVIRWIGQPFGLNEVLAGELEDEACGCTDGTFRGTRYFTCA
 LKALFKVLKSRPDSRFASLOPVSQIERCNSLAFGYLSVEVENTPPKMEKEGLE
 INMKKGIQGHVNSYDSTLPCLFAPSSVLDTLVRPKKNDVYYSSETOLLRLK
 IYNPLRIYGVCACTIMKILKILEVMAAGSTSEKPEEFNLILFPHILLRVEPLLK
 IRSAGKVQDCFPYQIFMSEKNEKVGUPTTQQLLEWFSFNSNLKFAEAPSCLLIQMRPF
 KDFKLFPKLIIPSLELIQYILKLPDSAGYVEGLQCMVENATTIRTSAGKIKQFC
 KTCNTQHLHPRLNHYNPVLKDLPMWRHGCIPQCNMELFVLICETSHYVAF
 VKYKDDSAWLFPDSMADRGGQNGFNIPQVTPCEVGYLXMSLEDLHSLDSRRIPR
 LCTKTAL"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3496
 Score: 4494.00 Matches: 917
 Percent Similarity: 95.33% Conservative: 2
 Best Local Similarity: 95.12% Mismatches: 30
 Query Match: 89.27% Indels: 22
 DB: 3 Gaps: 4

US-09-671-687A-3 (1-949) x AF161542 (1-3496)

QY 1 MetSerSerGlyLeuTrpSerGlnGlu-LysValThrSerProTyrTrp-GluGluArgI 20
 Db 248 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCACTCCCTACTGGGGGAAGAGCGGA 307
 QY 20 lePheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeu 40
 Db 308 TTTTCTTCTGCTTCTTCAAGAAATGAGCGGTACAGCAACAAACAAACAAAGCTCTCTTA 367
 QY 40 ysValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgI 60
 Db 368 AAGTACCGAA-GGAGTATAGGACAGTATATTCAGATCGCTC-GTGGGCAATCAAGGA 425
 QY 60 leProSerAlaLysGlyLysLeuAsnGlnIleGlyLeuLysIleLeuGluGlnProHisA 80
 Db 426 TTCCTTCTGCAAAAGGCAG-AAAAATCAGATTGGATTAAAAATTCATAGAGCAACCTCATG 484
 QY 80 laValLeuPheValAsp-Glu---AspValValGluIleAsnGlnLysPheThrGluLeu 98
 Db 485 CAGTCTCTTTGTGTGTAAGAAAGGATGTTGTAGAGATGAATGAAAGATTACAGAGATTT 544
 QY 99 LeuLeu---AlaIleThrAsnCys-GluGluArgPheSerLeuPheLysAsnArgAsnAr 117
 Db 545 ACTTTTGGGCAATACCCCAATTGTGGAGAGAGGTTTCAGGCTGTGTTAAAAACAGAAACAG 604
 QY 117 gLeuSerLysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerG1 137
 Db 605 ACTAAGTAAAGGCTCCCAATAGAGCGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGG 664
 QY 137 yGluGluLysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrVa 157
 Db 665 GGAAGAAAAATTTCTCGAGTTGTACGCTTTCAGAGGACCCCTGTGTAGCAGAGAGGACAGT 724
 QY 157 lSerGlyIlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAs 177
 Db 725 CTCGGGAATATCTTTTCGAGTTGAATTTGTG-GAAGAAGGTCGTGGTCAAGGTTTCACTGA 783
 QY 177 pGlyValTyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAl 196
 Db 784 CGGGGTGTACCAAGGGAAACAGCTTTTTCAGTTGTGATGAATATTGTGCGCGTGTGTTGTC 843

QY 196 aLeuAspLysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaG1 216
 Db 844 ATTGGCAAGCTAGAACTCATAGAGATGATGACACTGCATTGGAAAGTATTACGCAGG 903
 QY 216 yProGlyAspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLe 236
 Db 904 TCCTGGGGACACAATCAGTCGAACTTCCTCTCTTTGGAAATAAATCCCTCCAGAGTTCTTT 963
 QY 236 uLysGlyGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLy 256
 Db 964 GAA-GGTTGAGAAACAATAGACTCTGGACAGTTATATCTGTGATGTTTTGCCAGAAA 1022
 QY 256 sGluSerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspG1 276
 Db 1023 AGAAAGCTTAGGATATTTTGGTGGTGTGGACATGGATAACCTTATTGGCAACTGGGATGG 1082
 QY 276 yArgPheAspGlyValLeuCysSer-PheAlaCysValGluSerThrIleLeuLeuHisI 296
 Db 1083 AAGATTGTATGGAGCGCAGCTTGTAGTTTGCGTGTGTGAAAGTCCAATTTCTATTGCACA 1142
 QY 296 leAsnAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheM 316
 Db 1143 TCAATGATATCATCCAGAGAGTGTGACGAGGAAAGAGGCGCTTCCAAACTTGCCTTTTA 1202
 QY 316 etSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlyS 336
 Db 1203 TGTCAAGAGGTGTGGGACAAAGGTTTCATC-AGTCATATAAACAAGAGCTACAGGAT 1261
 QY 336 erThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerS 355
 Db 1262 CTACCTTAGACCTCGGAAATAGAAAACAGATCTGAATTTATTTATACCTTAAATGGGTCTT 1321
 QY 355 erValAspSerGln-ProGlnSerLysSerLysAsnThrTrpTyr-IleAspGlu-ValA 374
 Db 1322 CTGGTGACTTACAACTCCCAATCCAAATCAAAATCAAAATCCATGGTACCATTGATGAAAGTTG 1381
 QY 374 laGluAspProAlaLysSerLeuThrGluIleSerThr-AspPheAspArgSerSerPro 393
 Db 1382 CATAAAGACCTTGAAAAACTCTTACAGAGATATCTACCCGACTTTGGCCGGTCTTCACCA 1441
 QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
 Db 1442 CCACCTCCAGCTCTCTCTGTGAACCTCACTGACCACGAGAACAGATTCCACTCTTTACCA 1501
 QY 414 PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
 Db 1502 TTCAGTCTCACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTG 1561
 QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeu 453
 Db 1562 TCAGCCGAGTCTGTATATGGAAGAGCTAAACACTGCACCCGTCGAAGAGATCCACCTTG 1621
 QY 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
 Db 1622 GCATGCTCTCTGGAACTCACATGCTTAGAGTGGGCTCATTTGGTGAAGTTAAGGAG 1681
 QY 474 AsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
 Db 1682 AACCTCTCTTATGGGGTAATCCGTTGGATCGGTGAGTCCAGCCACCCAGGACTGAATGAAGTG 1741
 QY 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
 Db 1742 CTCGTGGACTGGAACCTGGAAGATGAGTGTGACGGCTGTACGGATGGAACCTTCAGAGGC 1801
 QY 514 ThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
 Db 1802 ACTCGGTATTTCACTGTGCCCTGAAGAAAGCGCTGTTTGTGAAACTGAAGAGTGCAGG 1861
 QY 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
 Db 1862 CCTGACTCTAGTGTGATTCATTTGCAGCCGCTTCCAAATCAGATTGAGCGCTGTAACCTCT 1921

554 LeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
1922 TTAGCATTTGGAGGCTCTTAAGTGAAGTGTAGAGAAATACTCCACCAAAATGGAA 1981
574 LysGluGlyLeuGluMetIleGlyLysLysGlyGlyIleGlnGlyHisTyrAsnSer 593
1982 AAAGAAGGCTTGGAGATAATGATTGGGAAGAAGAGGCATCCAGGTCATTACAATTCT 2041
594 CysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
2042 TGTACTTAGACTCAACCTTATTCTGCTTATTGCTTTAGTTCTGTTCTGGACACTGTG 2101
614 LeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 633
2102 TTACTTAGACCCCAAGAAAGAACGATGTAAGATATTATAGTGAACCCCAAGAGCTACTG 2161
634 ArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMet 653
2162 AGACAGAAATTTGTTAATCTCTGAGATATATGATATGTTGTGCCACAAAATATG 2221
654 LysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLys 673
2222 AAACCTGAGAAATACTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAA 2281
674 AspProGluGluPheLeuAsnIleLeuPheHisHisIleLeuArgValGluProLeu 693
2282 GATCTCGAGAAATCTTGAATATTTCTGTTTCATCATATTTAAGGGTAGAACCTTTGCTA 2341
694 LysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGlu 713
2342 AAATTAAGATCAGCAGTCAAGAGTACAGATTTTACTTCTATCAATTTTATGGAA 2401
714 LysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsn 733
2402 AAAAATGAGAAAGTGGCGTTCCCAATTCAGCAGTTGTTAGATGTTCTTTTATCAAC 2461
734 SerAsnLeuLysPheAlaGluAlaProSerCysIleuIleleGlnMetProArgPheGly 753
2462 AGTAACCTGAAATTTGAGAGGACCATCATATGCTGATTATTCAGATGCCCTCGATTGGA 2521
754 LysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeu 773
2522 AAAGACTTTAACTATTAAAAAAATTTTCCCTCTCTGGAAT-ATATACAGATTTA 2580
774 LeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArg 793
2581 CTTGAAGACACTCCGACAGTCCGGATATGTGGAGGCTTGCATGTATGATGATAGA 2640
794 GluCysTyrAspAspProAspIle-SerAlaGlyLysIleLysGlnPheCysLysThrCy 813
2641 GAATGCTACGACGATCCGACATTCCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTG 2700
813 sAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLe 833
2701 CAACACTCAAGTCCACCTTCATCCGAAGAGCGTGAATCAATAATATAACCCAGTGTCAC 2760
833 uProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetG1 853
2761 TCCCAAGACTTACCCGACTGGAGCTGAGACACCGCTGATCCCTTGCACAGATATGGA 2820
853 uLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLy 873
2821 GTTATTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGAA 2880
873 sAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnG1 893
2881 GGACGATTTCTGCTGCTCTCTTTTGACAGCATGGCCGATCGGGATGGTGGTCAAGATGG 2940
893 yPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLe 913
2941 CTTCAACATTCCTCAAGTCAACCCCATCCAGAGTAGGAGAGTACTTGAAGATGCTCT 3000
913 uGluAspLeuHisSerLeuAspSerArgArgIle-GlnGlyCysAlaArgArgLeuLeuC 933

3001 GGAAGACCTGCATCTCTTGGACTCCAGAGATCCCAAGGCTGTGCAGAAAGCTGCTTT 3060
933 ysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
3061 GTGATGCATATATGTGCTATGACAGAGTCCAAATGAGTTTGTACAAA 3110
RESULT 2
AY406374 2862 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens CYLD gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY406374
VERSION AY406374.1 GI:39762348
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2862)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene tritos
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2862)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source 1..2862
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>2862
/genes="CYLD"
/locus_tag="HCM2536"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2862
Score: 4313.00 Matches: 839
Percent Similarity: 88.04% Conservative: 0
Best Local Similarity: 88.04% Mismatches: 110
Query Match: 85.68% Indels: 4
DB: 9 Gaps: 4
US-09-671-687A-3 (1-949) x AY406374 (1-2862)
Qy 1 MetSerSerGlyLeuTyrProSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 1 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCTACTTCACCCCTCTGGAAGAGCGGATT 60
Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 61 TTTTACTTGTCTTCTCAAGATGCGAGCTTACAGACAAACAAACAAACAAACCTCTTAAA 120
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 121 GTACCGAAGGGAAGTAGGACAGTATATTCAAGATCGTTCTCTGTGGGCAITCAAGGATT 180
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 181 CCTTCTGAAAGCAAGAAAAATCAGATTGGATTAAAAATTTAGAGCAACCTCATGCA 240

QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 241 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGATTTCACAGAGTTACTT 300
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 301 TTGGCAATTTACCAATTTGTGAGGAGAGGTTGAGGCTGTTTAAACAGAAACAGACTAAGT 360
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 361 AAAGGCTCCAAATAGACGTGGGCTGCTGTGTGAAGATACAGCTGAGATCTGGGGAAGAA 420
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 421 AAATTTCTCTGGAGTTGTACGCTTCAGAGACCCCTGTTAGACAGAGGACAGTCTCCGGA 480
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
DB 481 ATATTTCTTTGGAGTTGAATTTGCTGGAGAAGGTCGTGTCAGGTTTTCACGTGACGGGGTG 540
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 541 TACCAAGGGAACAGCTTTTTCAGTGTGATGAGATTGGCGTGTGTTGTTGCTTTGGAC 600
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 601 AAGCTAGAACTCATAGAGATGATGACCTGCATTGGAAAGTGATTACGCGAGTCTCTGGG 660
QY 219 AspThrMetGlnValGlu---LeuProProLeuGluIleAsnSerArgValSerLeuLys 237
DB 661 GACACAATGCAGGTGCAANN 720
QY 238 GlyGlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGlu 257
DB 721 NNN 780
QY 258 SerLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArg 277
DB 781 NNN 840
QY 278 PheAspGlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 841 NNN 900
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 901 NNNNNNNNNNNNAGAGTGTGACGAGGAAAGAGGCTCCCAAACTTGCTTTATGTCA 960
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 961 AGAGGTGTTGGGACAAAGGTTTCATCCNGTCATAATAAACCAAGAGCTACAGATCTACC 1020
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1021 TCAGACCTCGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTCTGTT 1080
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1081 GACTCAACCAACCAATCCAAATCAAAATAATACATGTTGATGAAAGTTGACAGAGAC 1140
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1141 CTTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACTCCAG 1200
QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1201 CCTCTCTCTGAACTCACTGACCCAGAGAACAGATTCACACTCTTTTACCATTTCAGTCTC 1260
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1261 ACCAAGATGCCCAATCAACATGGAAGTATTGGCCACAGTCCACTTTCTCTGTCAGCCAC 1320

QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1321 TCTGTAAATGGAAGAGCTAAACACTGCACCCGTCCTCAAGAGAGTCCACCTTTGGCCATGCCT 1380
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1381 CTTGGGAATCTCATGCTGTCTAGAGTGGGCTCATTTGGCTGAAAGTTAAGGAGAACCTCTCT 1440
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1441 TTCTATGGGGTAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTCTCGCTGGA 1500
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1501 CTGGAACCTGGAAGATGATGTGAGGCTGTGACGATGTGACGAACTTCAGAGGCACCTCGGTAT 1560
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1561 TTCACTGTGCCCTGAAGAAGGCGCTGTTTGTGAACTGAAGAGCTGACGCTGACTCT 1620
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1621 AGGTTTGCATCATTCGACGCGGTTTCCAATCAGATTGAGCGCTGTAACCTCTTAGCATTT 1680
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 1681 GGAAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGGC 1740
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1741 TTGGAGATTAATGATTGGGAAGAAAGGCGCATCCAGGCTCATTAACAATCTTGTACTTA 1800
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 1801 GACTCAACCTTATCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTACTAGA 1860
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 1861 CCCAAAGAAAAGAACGATGTAGAAATATATAGTGAACCCCAAGAGCTACTGAGACAGAA 1920
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 1921 ATTGTTAATCTCTGAGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1980
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 1981 NNN 2040
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2041 GAATCTTGAATATCTGTTTCATCATATTTAAGGGTGAACCTTTGCTAAAAATAAGA 2100
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2101 TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG 2160
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2161 AAAGTTGGCGCTTCCCAATTCAGCAGTGTGTAAGATGGTCTTTTATCAACAGTAACCTG 2220
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
DB 2221 AAATTTGCAGAGGCACCATCATGTCTGATTATTTCAGATGCTCGATTTGGAAAAAGACTTT 2280
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2281 AAACTATTTAAAAAATTTTCTCTCTCGAAATTAATAATAACAGATTTTACTTTGAAGAC 2340
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2341 ACTCCACAGACAGTCCCGATATGTGGAGGGCTTCAATGATGATGATGATGAGATGCTTAC 2400
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816


```
Db 2401 GACGATCCGACATCTCAGCTGAAATAACAGCAGCTTTTGTAAACCTGCAACATCTCAA 2460
QY 817 ValHisLeuHisProLysArgLeuAenHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2461 GTCCACCTTCATCCGAAGAGGCTGAATCAATAATATAAACCAGTGTCACTTCCCAAAGAC 2520
QY 837 LeuProAspTTPAspTTPArgHisGlyCysValleProCysGlnAenMetGluLeuPheAla 856
Db 2521 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGAGATTATTGTCT 2580
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2581 GTTCTCTGATAGAAACAGACCACTATGTGTCTTTGTGAGATATGGAGAGCAGATTCT 2640
QY 877 AlaTTPLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAenGlyPheAenIle 896
Db 2641 GCCTGGCTCTCTTTCACAGCATGGCCGATCGGATGGTGGTTCAGAAATGGCTTCAACATT 2700
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2701 CCTCAAGTCACCCCATGCCAGAGTAGGAGATGACTTGAAGATGTCTCTGGAAGACCTG 2760
QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2761 CATTCTTGACTCCAGGAGATCCAAAGGCTGGCAGAGACTGCTTGTGTATGCATAT 2820
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 2821 ATGTGCATGTACCAGATCCAAACATGAGTTGTACAAA 2859

RESULT 3
AY406376
LOCUS
DEFINITION Mus musculus CYLD gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406376
VERSION 1
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2730)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2730)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source
location/Qualifiers
1..2730
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2730
/genes="CYLD"
/locus_tag="HMC2536"

gene
ORIGIN
Alignment Scores:
```

```
Pred. No.: 0 Length: 2730
Score: 4102.00 Matches: 794
Percent Similarity: 85.50% Conservative: 20
Best Local Similarity: 83.40% Mismatches: 92
Query Match: 81.49% Indels: 46
DB: 9 Gaps: 5

US-09-671-687A-3 (1-949) x AY406376 (1-2730)
QY 1 MetSerSerGlyLeuTTPSerGlnGluLysValThrSerProTyrTTPGluGluArgIle 20
Db 1 ATGAGTTTCAGGCTGTGGAGCCCAAGAGAAAGTTACTTTCACCTACTCGGAAGAACGGATT 60
QY 21 PheTyrLeuLeuLeuGlnCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db 61 TTTTATCTGCTTCTTCAAGAAATCAGTGTAACTAACACAACTCAAGAGCTCTCTGAAA 120
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 121 GTACCCAAAGGAGTATAGGACAGTATCATCAAGACCTTCTGTGGGCGATTCAAGAGTT 180
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 181 CCTTCCACAAAAGCAAGAAAATCAGATTGGATTAAATAATCTTGGAGCAACCCGATGCA 240
QY 81 ValLeuPheValAspGlu---AspValValGluLleAsnGluLysPheThrGluLeuLeu 99
Db 241 GTTCTCTGTTTTCATGAAAAAGGATGTTGTAGAAAATAAATGAAAAATTCACAGAGTTACTG 300
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 301 TTGGCAATTACCACACTGTGTAGGAGAGGCTCAGGCTATTATAGAAACACAGACTCCGAGT 360
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 361 AAGGGCTCCAGGTAGACGTGGGCGAGTCTGTGAAAGTACAGTCCGATCTGGGGAGAG 420
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 421 AATTTTCCAGGAGTGTGTAGCTTTCAGAGGACCTTTATTAGCGGAGAGGAGCGTGTCCGG 480
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
Db 481 ATTTTCTTTGGAGTAGAAATTTATTGGAAGAGGCTGTGTCAAGGTTTTCACGGATGGGGTA 540
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 541 TATCAAGGGGAAGCAGCTTTTCCAGTGTGATGAGAGCTGTGGCGTGTGTTTGTTCATTGGAC 600
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 601 AACTAGAACTTATAGAGATGATGACAATGGATTGGAAGATGATTTTCAGGCGCCAGGA 660
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 661 GATACAATGCAGGTTGAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 721 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 744
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAenTTPAspGlyArgPhe 278
Db 744 ----- 744
QY 279 AspGlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAsp 298
Db 745 ----- 745
QY 299 IleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArg 318
Db 775 NNNNNNNNNNATAGCGTGACACAGGAAGAGGCGCTCCCAAACTTGCCTTTATGTCAAGA 834
```

QY 319 GlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSer 338
 Db 835 CGTGTAGGTGACAAAGGTTTCATCTAGTCATAATAACCAAGGTTACAGGATCTACCTCA 894
 QY 339 AspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAsp 357
 Db 895 GACCCCTGGGAAGTAGAAACAGATCTGAATATATTTATACCTTAAATGGGTCATCTGTGAC 954
 QY 358 SerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspPro 377
 Db 955 TCACAA--CAATCCAAAGTCCAAAATCCATGGTACATTTGATGAAGTTGCAGAGACCT 1011
 QY 378 AlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProLeuGlnPro 397
 Db 1012 GCAAAGTCACTTACAGAGATGCTTCGGACTTCGGACATTCATCTCTCCACCGCAGCCT 1071
 QY 398 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThr 417
 Db 1072 CCTTCCATGAACCTCTGTCTAGCGAAGACAGATTCCATCTCCCTTACCCCTTACGCTGACA 1131
 QY 418 LysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSer 437
 Db 1132 AAGATCCCAATACTAATGCACATGGCTCATAGTCCATCTCTCTGTCAGTGCAGTCT 1191
 QY 438 ValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeuAlaMetProPro 457
 Db 1192 GTGATGGGGAGCTGAACGACACCTGTCCAGGAGTCCACCTTGCCCATCTCTTCT 1251
 QY 458 GlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPhe 477
 Db 1252 GGGAAATGCACACCGGCTAGAGGTGGGCTCAGTGGCTGAAGTAAAGAGAACCCTCCCGGTT 1311
 QY 478 TyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeu 497
 Db 1312 TATGGGGTTATCCGTTGGATGGCCAGCCACCGAGGCTCAGTACGCTGAGTGTGAGT 1371
 QY 498 GluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPhe 517
 Db 1372 GAACCTGGAAGATGAATCGCAGGCTGTACAGATGGAACITTCAGGGGCGACCGGTATTTC 1431
 QY 518 ThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArg 537
 Db 1432 ACGTGTGCCCTGAAGAAGGCACTGTTTGTGAACCTGAAGAGCTGCAGACCGGACTCTAGG 1491
 QY 538 PheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGly 557
 Db 1492 TTGGCATCTTGCAGGCTGTTTCCATCAAAATGAAGGGTGAACCTTTTAGCATTTGGG 1551
 QY 558 GlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeu 577
 Db 1552 GGTATTTTAAGTGAAGTAGTAGAAGAAATATCTCCACCTAAAATGGAAAGGAGGTTTA 1611
 QY 578 GluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAsp 597
 Db 1612 GAGATAATGATTGGAAAGAAAGGCAATCCAGGGCCATACAAATCTTGTTACTTAGAC 1671
 QY 598 SerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgPro 617
 Db 1672 TCAACTTTATCTGCTTATTTGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1731
 QY 618 LysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIle 637
 Db 1732 AAAGAGAGAGATGATAGAGTATTACAGTGAGACTCAGAGGCTACTGAGACAGAGATA 1791
 QY 638 ValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLys 657
 Db 1792 GTCAATCTCTGAGATNN 1851
 QY 658 IleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGlu 677
 Db 1852 NNN 1911
 QY 678 PheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSer 697

Db 1912 TTCTTAATATCTGTTTCATGATATTTTAAGGTTGAACCATTTGTTAAAAATAAGATCA 1971
 QY 698 AlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLys 717
 Db 1972 GCAGGTCAAAAGTTCCAGACTGTAACTCTCTATCAAAATTTTATGAAAAAATGAGAA 2031
 QY 718 ValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLys 737
 Db 2032 GTTGGAGTACCACAAATTCAGCAGTTATTAGAAATGGTCTTTTATCAACAGCAACCTGAAA 2091
 QY 738 PheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLys 757
 Db 2092 TTTGCAGAGGCACCATCATCTTGATTATCCAGATGCTCGTTTGGAAAAAGACTTTAAA 2151
 QY 758 LeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThr 777
 Db 2152 CTATTTAAAAAAATTTTCTTCCCTGGAAATTAATAATAACAGATTTTACTTTGAAGACCT 2211
 QY 778 ProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyrAsp 797
 Db 2212 CCCAGGAGTCCGCATCTGTGGAGACTTGCATGTACGAGTGCAGGAGTGTCTATGAC 2271
 QY 798 AspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnVal 817
 Db 2272 GATCCGGACATCTCAGCTGGAAAGATCAAGCAGTTCTGTAAAGACCTGCAGACCTCAGGTT 2331
 QY 818 HisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeu 837
 Db 2332 CACCTTCATCCCGAGGTTGAATCATTTATCATCCAGTATCATTCTCCCAAGACTTG 2391
 QY 838 ProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaVal 857
 Db 2392 CTTGACTGGAGTGGAGACACGCTGCATCCCTGTGCAGAGATGGAGTTATTGCTGTT 2451
 QY 858 LeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSerAla 877
 Db 2452 CTCTCATAGAAACGAGCCACTATGTTGCTTTGTGAAGTATGGAAAGGATGACTCTGCC 2511
 QY 878 TrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIlePro 897
 Db 2512 TGGCTTTTCTTTGACAGCATGCGGATCGAGATGGTGGTTCAGAAATGGCTTCAACATTCCA 2571
 QY 898 GlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHis 917
 Db 2572 CAAGTACGCCCTGCCAGAGTGGGAGAGTACTTGAAGATGCTCTCTGGAGGACCTGCAC 2631
 QY 918 SerLeuAspSerArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyrMet 937
 Db 2632 TCTTTGGATTCCAGAGGATTCAAGGCTGTGCGCGCAGACTTCTTTGCGATGCATACATG 2691
 QY 938 CysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 2692 TGCATGTACCAGATGCCAACCATGAGCCTGTACAAA 2727
 RESULT 4
 AY406375 2862 bp DNA linear GSS 15-DEC-2003
 LOCUS Pan troglodytes CYLD gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY406375
 VERSION AY406375.1 GI:39762349
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1 (bases 1 to 2862)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
 Todd,M.A., Fanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 2862)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, P., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES
 source
 1..2862
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>2862
 /gene="CYLD"
 /locus_tag="HCW2536"

ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2862
 Score: 3768.00 Matches: 751
 Percent Similarity: 78.80% Conservative: 0
 Best Local Similarity: 78.80% Mismatches: 198
 Query Match: 74.85% Indels: 4
 DB: 9 Gaps: 4

US-09-671-687a-3 (1-949) x AY406375 (1-2862)

QY 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTyrGluGluArgIle 20
 |||||
 DB 1 ATGAGTTTACGGCTTATGGAGCCAGAAAGTCACTTCCCTTACTTGGGAAGCGGATT 60
 |||||
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 |||||
 DB 61 TTTTACTTGTCTTCAAGATGTCAGCGTTACAGACAAACAAACAAAGCTCTTAAA 120
 |||||
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 |||||
 DB 121 GTACCGAAGGGAAGTATAGGACAGTATATTCAGATCGTTCGTGGGGCATTTCAAGGATT 180
 |||||
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 |||||
 DB 181 CCTTCTGCAAAAGGCAAGAAATCATGATTGGATTAAAAATTTAGAGCAACCTCATGCA 240
 |||||
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 |||||
 DB 241 GTTCTCTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGATTACTT 300
 |||||
 QY 100 LeuAlaIleThrAsnGlyGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 |||||
 DB 301 TTGGCAATATACCAATTNNNNNAGAGTTTCAGCCTGTTTAAAAACAGAAACAGACTAAGT 360
 |||||
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 |||||
 DB 361 AAANNNTTCCAAATAGNN 420
 |||||
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 |||||
 DB 421 NNN 480
 |||||
 QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 |||||
 DB 481 NNN 540
 |||||
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 |||||
 DB 541 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTGTCATTGGAC 600
 |||||

QY 199 LysLeuGluLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 |||||
 DB 601 AAGCTAGAACTCATAGAAATGATGACACATTTGAAATGATTAACGCGAGGCTCTGGG 660
 |||||
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
 |||||
 DB 661 GACACATGCGAGTTCGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720
 |||||
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 |||||
 DB 721 NNNNNNACAATAGAACTCGAAACNNNTATATTTCTGTGATGTTNNNNNNNGAAAGAAAGC 780
 |||||
 QY 259 LeuGlyTyrPheValGly---ValAspMetAspAsnProIleGlyAsnTyrAspGlyArg 277
 |||||
 DB 781 TTAGGATATTTTGTGTGTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 840
 |||||
 QY 278 PheAspGlyValLeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 |||||
 DB 841 NNN 900
 |||||
 QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
 |||||
 DB 901 NNNNNNNNNNNNNNNNTGTGCGCAGAAAGGAGGCTTCCCAAACTTGCCTTTATGTCA 960
 |||||
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 |||||
 DB 961 AGAGGTGTTGGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGTTACAGGATCTACC 1020
 |||||
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerVal 356
 |||||
 DB 1021 TCAGACCTTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTTCTGTT 1080
 |||||
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
 |||||
 DB 1081 GACTCACACCACAATCCAAATCAAAAATACATGGTACATTTGATGAAGTTGCAGAACAC 1140
 |||||
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
 |||||
 DB 1141 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTACCGTTCTTACCACCACTCCAG 1200
 |||||
 QY 397 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 |||||
 DB 1201 CCTCTCTCTGTGAACCTACTGACCCGAGAGACAGATTCCACTCTTTACCATTGAGTCTC 1260
 |||||
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 |||||
 DB 1261 ACCAAGATGCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTTCTGTGAGCCAG 1320
 |||||
 QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 |||||
 DB 1321 TCTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCCCTTGGCCATGCT 1380
 |||||
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 |||||
 DB 1381 CCTGGGAACCTCACATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCT 1440
 |||||
 QY 477 PheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 |||||
 DB 1441 TTTATGGAGTATCTCGTGGATCGGTGCGTCAGCCACAGACTGAATGAAGTCTGCTCGA 1500
 |||||
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 |||||
 DB 1501 CTGGAACCTGGAAGATGAGTGTGCGAGGCTGTACGGATGGAACTTTCAGAGGCACCTCGGTAT 1560
 |||||
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
 |||||
 DB 1561 TTTCACTGTGCTTGAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGTCAGGCTGACTCT 1620
 |||||
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 |||||
 DB 1621 AGTTTTCATCATTTGACCGCGTTTCCAATCAGATCGAGCGCTGTAATCTCTTTAGCATTT 1680
 |||||
 QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
 |||||

Db 1681 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAGAGGC 1740
Qy 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1741 TTGGAGATAATGATTGGGAAGAAGAACGATCCAGGGTCATTACAACTCTTGTTACTTA 1800
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 1801 GACTCAACCTTATCTGCTATTGCTTTAGTTCTGTGACACTGTGTTACTTAGA 1860
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 1861 CCCAAGAAAAAGAACGATGAGAATATTATAGTGAACCCAGAGACTCTGAGGACAGAA 1920
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 1921 ATTGTTAATCTCTGAGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1980
Qy 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 1981 NNN 2040
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2041 NNN 2100
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2101 TCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGGAAGAAAATAG 2160
Qy 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2161 AAAGTTGGCTGCCACATTCAGCAGTGTGTAGATGGTCTTTATCAACAGTAACCTG 2220
Qy 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
Db 2221 AAATTTGCAGAGNN 2280
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAsp 776
Db 2281 NNN 2340
Qy 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2341 NCTCCAGACAGTCCGGATATGTGAGGGCTTGCAATGTATGAGTGTAGAGATGCTAC 2400
Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2401 GACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTTGTAAACCTGCAACACTCA 2460
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2461 GTCCACCTTCATCCGAAGAGGCTGAATCAATAATAAACCCAGTGCATCTCCCAAGAC 2520
Qy 837 LeuProAspThrAspThrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2521 TTACCTTGACTGGAGCTGGAGACAGCGTGCATCCCTTGCAGAAATATGGAGTTATTGCT 2580
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2581 GTTCTCTGCATAGAAACCAACGCACTATGTTGCTTTGTGAAGTATGGGAAGACGATTCC 2640
Qy 877 AlaThrLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2641 GCCTGGCTCTCTTTGACAGCATGGCGCATCGGATGGTGGTGCAGAACCGCTTCAACATT 2700
Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2701 CCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGACCTG 2760
Qy 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936

Db 2761 CATTCCTTGGACTCCAGGAGATCCAGGCTGTGCAGAGACTGCTTTGTGATGCATAT 2820
Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrIlys 949
Db 2821 ATGTGCATGTACCAGAGTCCAAACATGAGTTTGTACAAA 2859
RESULT 5
AK039054
LOCUS
DEFINITION
Mus musculus adult male hypothalamus cDNA, RIKEN full-length
enriched library, clone:A230090N11 product:CYLINDROMATOSIS (TURBAN
TUMOR SYNDROME) homolog (Homo sapiens), full insert sequence.
AK039054
VERSION AK039054.1 GI:26333008
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P., and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
6 (bases 1 to 2220)
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
7 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

FEATURES

source

1. .2220
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A230090N11"
/db_xref="taxon:10090"
/clone="A230090N11"
/sex="male"
/tissue_type="hypothalamus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
352. .>2220

CDS

/note="unnamed protein product; CYLINDROMATOSIS (TURBAN
TUMOR SYNDROME) homolog [Homo sapiens] (SPTR1Q96SH0,
evidence: FASTV, 93.8%ID, 30.5%length, match=873)
putative"
/codon_start=1
/protein_id="BAC30222.1"
/db_xref="GI:26333009"
/translation="MSSGLWSQEKVTSPPYWEIRIFYLLEQESVTDKQKLLKPKVG
STGVYIDRSVGHSPVSKGKNOIGLILQPHAVLFVDEKDVVEINEKFTLLLA
TINCERLSLRNRLSKGLQVDSQPKVQLRSGEKFPVGRGPIILARTVSG
IFPVELLEEGQGTDFGVYQKQLFQDCEDGVFVLDKLELIEDDNGLESDFAG
PDGTWQVEPPLIENSRSLKVGESTETVIFCDVLPKESLGYFVGVDMNPIGNW
DRFPGVQLCSFASVSTILLINDIIPALSDSVTOERRPPKLAFFMRGVGDKGSSSH
NPKVTSPPQPSVNSRSLFVTLNGSSVDSQSKNPWIDEVAEDPAKSLTMS
DFGHSPPQPSVNSRSLFVTLNGSSVDSQSKNPWIDEVAEDPAKSLTMS
TPVQSPPLPISGNHGLEVSLAEVKNPFFYVIRWIGQPPGLSDVLGLELDE
CAGCTDGTFRGTFTYFCAUKALFVKLSRPSRFSALQVSNQIERNCSLAFGGYL
SEVVEPTPPRMEKEGLEIMIGKKGIQGHYNSCYLDSTLFLCLFAPSSALDVTLLRP"

ORIGIN

Alignment Scores:
Pred. No.: 8,73e-294 Length: 2220
Score: 2952.00 Matches: 575
Percent Similarity: 94.87% Conservative: 17
Best Local Similarity: 92.15% Mismatches: 24
Query Match: 58.64% Indels: 8
DB: 3 Gaps: 6

US-09-671-687A-3 (1-949) x AK039054 (1-2220)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 352 ATGAGTTTACGGCTGTGGAGCCCAAGAGAAATGTTACTTCCCTACTGGGAGAACCGATT 411
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 412 TTTTATCTGCTCTTCAAGATGCAGTGTAAACAGACAAACAACTCAGAGCTGCTGAA 471
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 472 GTACCCAAAGGGAGTAGACAGTACATCCAAAGCCGTCTCTGTGGGGCATTCAGAGTT 531
QY 61 ProSerAlaLysGlyLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
DB 532 CCTTCCACAAAGGCAAGAAATCAGATTGGATTAAATAATCTTGGGCAACCGCATGCA 591
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99

DB 592 GTTCTGTTGTTGATGAAGAGGATGTTGTAGAAATAAATGAAAATTCACAGAGTTACTG 651
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 652 TTGGCAATTAACCAACTGTGAGGAGAGGCTCAGCCTATTAGAAAACAGACTCCGACTAAGT 711
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 712 AAAGGGCTCCAGGTGACGCTGGCGAGCTCTGTGAAAGTACAGCTGCGCATCTGGGGAAG 771
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 772 AAATTTCCAGGAGTTGTACGCTTCAGAGGACCTTTATTAGCGGAGAGGACGGTGTGGGG 831
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 832 ATTTTCTTTGGAGTAGAATATTGGAAGAGGTCGTGCTCAAGGTTTTCAGCGATGGGTA 891
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 892 TATCAAGGGAAGCAGCTTTTCCAGTGTGATGAGGACTGTGGCGTGTGTTGTCATTGGAC 951
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 952 AAATAGAACTTATAGAAAGATGACAAATGGATTGGAAGTGAATTTTTCAGAGGCCAGGA 1011
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1012 GATACAATGCAAGTTGAACCTCCCTTTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT 1071
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1072 GGAGAAGTACAGAACTCTGGAACAGTAATATCTGTGATGTTTACCAGGAAAAGAGAGT 1131
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1132 CTAGGATATTTTGTGTGTGGACATGGAATACCTATTGGCAACTGGGATGGAGGTTT 1191
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1192 GATGGAGTACAGCTCTGTAGTTTTCAGAGTGTGAAAGTACAAATCTCTGCACATCAAT 1251
QY 298 AspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAla 314
DB 1252 GACATCATCCAGCTTTATCAGATAGCTGACACAGGAAAGAGGCGCTCCCAACTTGC 1311
QY 315 PheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThr 334
DB 1312 TTTATGTCAAGAGGTGTAGGTGACAAAGGTTTCATCTAGTCATAATAAACCAGAGTTACA 1371
QY 335 GlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGly 353
DB 1372 GGATCTACCTCAGACCTCGAAGTAGAAGACAGATCTGAAATATTATTATACCTTAATGG 1431
QY 354 SerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVal 373
DB 1432 TCATCTGTTGACTCACAA---CAATCCAAGTCCAAAAATCCATGTCATGATGAGAGTT 1488
QY 374 AlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPro 393
DB 1489 GCAGAAAGACCTCGCAAGTCACTTACAGAGATGCTTTCGGGACTTTCGGACATTCATCTCT 1548
QY 394 ProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
DB 1549 CCACCGCAGCTCTCTCCGTGAACCTCTGTGTCAGAGAGAACAGATTCATCTCTTACCC 1608
QY 414 PheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeu 433
DB 1609 TTAGCCTGCAAGATGCCCATACTAATGGCAGCATGGCTCATAGTCCACTCTCTCTG 1668
QY 434 SerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeu 453

```

Db 1669 TCAGTGCAGTCTGTGATGGGGAGCTGAACACACACCTGTCTCAGAGAGTCCACCCCTTG 1728
Qy 454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db 1729 CCCATCTCTTGGGAATGCACACGGCTAGAGTGGCTCACTGGCTGAAGTAAAGAG 1788
Qy 474 AsnProProPheTyGlyValLleArgTrpLleGlyGlnProProGlyLeuAsnGluVal 493
Db 1789 AACCCCGCTTCTATGGGCTTATCCGTTGGATTGGCCAGCCACAGGGCTCAGTGACGTG 1848
Qy 494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db 1849 CTAGCTGGACTGGAATGGAATGAATGCGCAGGCTGTACAGATGGAATTTTCAGGGGC 1908
Qy 514 ThrArgTyPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
Db 1909 AGCGGATTTTACGTGTGCTTGAAGAGGACACTGTTGTGAACTGAAGAGCTGCAGA 1968
Qy 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db 1969 CGGAGCTCTAGTTTGCATCTTTCAGCTGTTCATCAATCAAAATTGAAGGTGTAACCT 2028
Qy 554 LeuAlaPheGlyClyTyLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
Db 2029 TTAGCATTTGGGGCTATTAAAGTAGTAGTAGAAGAAATACCTCCACTAAATGGAA 2088
Qy 574 LysGluGlyLeuGluLleMetIleGlyLysLysLysGlyIleGlnGlyHisTyAsnSer 593
Db 2089 AAGGAGGTTTAGAGATANTGATTGGAAAGAGAAGGCAATCAGGGCCATACAACTTCT 2148
Qy 594 CysTyLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVal 613
Db 2149 TGTACTTAGACTCAACTTATTCTGCTATTGCTTTGCTTTGCTTCTGCTGAGCACTGTG 2208
Qy 614 LeuLeuArgPro 617
Db 2209 TTACTTAGACCC 2220

RESULT 6
LOCUS BC028885
DEFINITION Mus musculus cylindromatosis (curban tumor syndrome), mRNA (CDNA
ACCESSION BC028885
VERSION BC028885.1 GI:23468224
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S.,
Carninci P., Frange C., Raja S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S.,
Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,
Boutard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalios D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257

```

```

PUBMED 12477932
REFERENCE 2 (bases 1 to 3137)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 40 Row: g Column: 11
 This clone has the following problem: retained intron.

FEATURES

source
 1. .3137
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3965911"
 /tissue_type="Mammary tumor. Metallothionien-TGF alpha
 model. 10 month old virgin mouse. Taken by biopsy."
 /clone_lib="NCI CGAP Mam1"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:
 Pred. No.: 1.42e-274 Length: 3137
 Score: 2769.00 Matches: 546
 Percent Similarity: 66.00% Conservative: 15
 Best Local Similarity: 64.24% Mismatches: 16
 Query Match: 55.01% Indels: 273
 DB: 3 Gaps: 1

US-09-671-687A-3 (1-949) x BC028885 (1-3137)

Qy 373 ValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSer 392
 Db 186 GTTGCAGAGAGACCTCGAAAGTCACCTACAGAGATGCTTCGGACTTCGGACATTCATCT 245
 Qy 393 ProProLeuGlnProProValAsnSerLeuThrGluAsnArgPheHisSerLeu 412
 Db 246 CCTCCACCGCAGCTCTCTCCATGAACTCTTGTCTAGCGAGAACACAGATTCCTCTTA 305
 Qy 413 ProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSer 432
 Db 306 CCCTTCAGCTGACAAAGATGCCCAATACCTAATGGCAGCATGGCTCATAGTCCACTCTCT 365
 Qy 433 LeuSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProPro 452
 Db 366 CTGTCTAGTGCAGTCTGTGATGGGGAGCTGAACAGCACACTGTCTCAGGAGAGTCCACCC 425
 Qy 453 LeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLys 472
 Db 426 TTGCCCATCTCTTCTGGGAATGCACACGGGCTAGAGGTGGGCTCACTGGCTGAAGTAAA 485
 Qy 473 GluAsnProProPheTyGlyValLleArgTrpLleGlyGlnProProGlyLeuAsnGlu 492

Db 486 GAGAACCCCGTCTCTATGGGTTATCCGTTGGATTGGCCAGCCACCAGGGCTCAGTGAC 545
Qy 493 ValLeuAlaGlyLeuGluLeu----- 499
Db 546 GTGCTAGCTGGACTGGAACCTGTAATTTAATTTGGCCAAAAAATCTACAACCTCTTTTCCTG 605
Qy 499 ----- 499
Db 606 CATGAGGTTTGTCTAGATTCCCTCTATTAGATAATTTCTAGAAAAATCATCATGTGTG 665
Qy 499 ----- 499
Db 666 TGTGTGTGTATGTGTGCGTGATGTGTGTATTTCCAGTAAATAGAGAAAAATATTCTT 725
Qy 499 ----- 499
Db 726 ACTGTATGGCCTAGCTTCATTCACCCCTTTGCCATTCCTTAAACTTCTGGCTCATTTTTTC 785
Qy 499 ----- 499
Db 786 CTCGTGTTTCTGTCTGTAATCAAGAGTAGAGGAATTTTATAGGCTAGACTATTATAGCT 845
Qy 499 ----- 499
Db 846 TCCTTATATTAGTATAATGCACCTAACAAAGCAATTATCAGTTTTTCTTCTTACTCCTC 905
Qy 499 ----- 499
Db 906 AGCCTCCTCTGATGATTGGCTTTTACAAATTAATCATCATGTTCAGAGGCTTTTGGTAGA 965
Qy 499 ----- 499
Db 966 AGAGGTTTCCCTTTGGATATTAGAAAAAGAAATATAGTGTGATTTTCTATAATTATGAATG 1025
Qy 499 ----- 499
Db 1026 AGGAACCTGAATAAACCTTAGCTTTTACCTTAATGTGCTTTTTCATTAAATTTTTTTCATTT 1085
Qy 499 ----- 499
Db 1086 GTTATCTTGATGTGAGGAACCTAGATAAATTAATCTACCTTCCCGCCCTTGATCTTA 1145
Qy 499 ----- 499
Db 1146 ATACTTTTGAATGGCTCAGGGCAAAATCAGTAAACCAAGAAATGCATGAAGTACTTCCT 1205
Qy 499 ----- 499
Db 1206 GGTGGGCTAAAGAAATGTCTTCTTAAACCCTACCTCGTCAGCTAGTCAGCTCCCTTTTCT 1265
Qy 499 ----- 499
Db 1266 TGATGCTCTTCTATTTCAGCACAGATTGATGAAGGAATTCAGACAAAGTGTAAAGTCTTT 1325
Qy 500 -----G1 500
Db 1326 CCCCCTCTCTGCAAGGTGCAGTGTAACACATACCGAAGAAGACTTTCTTTTTCAGGA 1385
Qy 500 uAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyPheThrCysAl 520
Db 1386 AGATGAATCGCAGGCTGTACAGATGGAACCTTTTCAGGGGCACGCGGTATTTTCACGTGTC 1445
Qy 520 aLeuIysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSe 540
Db 1446 CTTGAAGAAGGCACCTGTTTGTGAACCTGAAGAGCTGCAGCCGACTCTAGGTTTGCATC 1505
Qy 540 rLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLe 560
Db 1506 CTTGCAGCCTGTTTCCAAATCAGATTGAAGGTCTAACTCTTAGCATTTGGGGGCTATT 1565
Qy 560 uSerGluValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluLeMe 580
Db 1566 AAGTGAAGTAGTAGAAAAAATACTCCACCCTAAATGGAAGGAAGGTTTAGAGATAAT 1625

Qy 580 tIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrIle 600
Db 1626 GATTGGAAAAAGAAAGGCATCCAGGGCCATTACAATTCCTTGTACTTAGACTCAACCTT 1685
Qy 600 uPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArgProLysGlyLys 620
Db 1686 ATTCTGCTTATTTGCTTTTATGTTCTGCCCTGGACACTGTGTGTACTTAGACCCCAAGANA 1745
Qy 620 sAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPr 640
Db 1746 GAATGATATAGAGTATTACAGTGAGACTCAGGAGCTACTGAGGACAGAGATAGTCAATCC 1805
Qy 640 oLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGly 660
Db 1806 TCTGAGAAATATATGGATATGTGTGCCACAAAGATTATGAACCTGAGGAAAAATACTTGA 1865
Qy 660 uLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAs 680
Db 1866 AAAAGTTGAGCTGCATCAGGTTTTTACTCTGGAAGAAAAAGATCCTGAAGAAATTTCTTAA 1925
Qy 680 nIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGly 700
Db 1926 TATCCTGTTTCATGATATTTTAAAGGTTGAACCATTTGTTAAAAATAGATCAGCAGGTCA 1985
Qy 700 nLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyVa 720
Db 1986 AAAAGTTCAAGACTGTAACTTCTATCAAAATTTTATGMAAAAAAATGAGAAAGTTGGAGT 2045
Qy 720 lProThrIleGlnLeuLeuGluTyrPheThrPheIleAsnSerAsnLeuLysPheAlaGly 740
Db 2046 ACCCAAAATTCAGCAGTATTAGAAATGTTTATCAACAGCAACCTGAAATTTGCAGA 2105
Qy 740 uAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760
Db 2106 GGCACCATCATGCTTGATATCCAGATCGCTCGGTTTGGAAAAAGACTTTAACTATTATA 2165
Qy 760 sLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGly 780
Db 2166 AAAAAATTTTCTTCCCTGGAATTAATAATAACAGATTTACTTGAAGACACTCCAGGCA 2225
Qy 780 nCysArgIleCysGlyLysLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAs 800
Db 2226 GTCCCGCATCTGTGGAGGACTTGCAGTGTACGAGTGCAGGGAGTGCTATGACGATCCGGA 2285
Qy 800 pIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHi 820
Db 2286 CATCTCAGCTGGGAAGATCAAGCAGTTCTGTAAAGCTCTGCAGCACTCAGGTTCACTTCA 2345
Qy 820 sProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTyr 840
Db 2346 TCCAGGAGGTTGAATCATTTCTTATCATCCAGTATCACTTCCCAAGACTTGCCTGACTG 2405
Qy 840 pAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysAl 860
Db 2406 GGACTGGAGACACGGCTGCATCCCTGTCAAGATGGAGTTATTGCTGTCTCTGTCAT 2465
Qy 860 eGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPh 880
Db 2466 AGAAACGAGCCACTATGTCTTTTGTGAAGTATGGGAAGGATGACTCTGCTGGCTTTT 2525
Qy 880 ePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValTh 900
Db 2526 CTTTGACAGATCGCGGATCGAGATGGTGTGTCAGATGGCTTCAACATTCACCAAGTGAC 2585
Qy 900 rProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAs 920
Db 2586 GCCTGCCCCAAGAGTGGGAGAGTACTTGAAGATGTCTCTGGAGGACCTGCACTCTTTTGA 2645
Qy 920 pSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTy 940
Db 2646 CTCCAGAAGGATTCAGAGGCTGTGGCGCAGACTTCTTTTGCATCATACATGTGTCATGTA 2705


```

QY 940 rClnSerProThrMetSerLeuTyrIlys 949
DB 2706 CCAGAGTCCACCATGAGCGCTGACAAA 2733

RESULT 7
BO433523 1063 bp mRNA linear EST 24-MAY-2002
LOCUS AGENCOURT_7766644 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015509
DEFINITION 5', mRNA sequence.
ACCESSION BO433523
VERSION BO433523.1 GI:21172599
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13211 row: 1 column: 06
High quality sequence stop: 672.

FEATURES
source
1..1063
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6015509"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1,97e-151 Length: 1063
Score: 1572.50 Matches: 317
Percent Similarity: 95.27% Conservative: 5
Best Local Similarity: 93.79% Mismatches: 12
Query Match: 31.24% Indels: 7
DB: 5 Gaps: 0

US-09-671-687A-3 (1-949) x BO433523 (1-1063)

QY 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuIlysIysAlaLeuPheValIysLeu 529
DB 1 ACCTTCAGAGGCGACTCGGTATTTTACCTGTGCGCTGAGGAAGCGCTGTTGTGAAACTG 60

QY 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerIleGlnIleGlu 549
DB 61 AGAGCTGAGGCGCTGACTCTAGTTTGCATCATTCAGCGCGGTTCCTCAATCAGATTGAG 120

QY 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrPro 569
DB 121 CGCTGTAACTCTTTAGCAATTTGAGGCTACTTAAGTGAAGTACTAGAGAATACTCCA 180

QY 570 ProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGly 589
DB 181 CCAGAAATGGAAGAAAGCGTTGGAGATAATGTTGGGAAGGAAGGCGATCCAGGGT 240

```

```

QY 590 HisTyrAsnSerCysTyrIleuAspSerThrLeuPheCysLeuPheAlaPheSerSerVal 609
DB 241 CATTACAATTTCTGTACTTAGACTCAACCTTATTCTGCTTATTGCTTTTAGTCTGTT 300

QY 610 LeuAspThrValIleuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThr 629
DB 301 CTGGACACTGTGTACTTAGACCCAAAGAAAGACGATGTAGATATATTAGTGAAC 360

QY 630 GlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAla 649
DB 361 CAAGAGCTACTGAGGACAGAAATTTGTAATCTCTGAGATATATGATATGTGTGCC 420

QY 650 ThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThr 669
DB 421 ACAAAATTTATGAACCTGAGGAAATCTTTGAAAGGTGAGGCTGCATCAGGATTTACC 480

QY 670 SerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgVal 689
DB 481 TCTGAAGAAAGAAAGATCTGAGGAATTTCTTGAATATTCTGTTTTCATCATATTTAAGGGA 540

QY 690 GluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGln 709
DB 541 GAACCTTTGCTAAATAATAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAA 600

QY 710 IlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrp 729
DB 601 ATTTTATGMAAAATGAGAAAGTTGCGCTTCCACAAATTCAGCAGTGTGTTAGAATGG 660

QY 730 SerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMet 749
DB 661 TCTTTTATCAACAGATGAACCTGAAATTTGACAGAGCCACCATCATGTCTGATTATTCAGATG 720

QY 750 ProArgPheGlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsn 769
DB 721 CCTCGAATTTTGAAGACCTTTAACTATTTAAAATTT-TTCTCTCTCTGGG-ATTAAAT 778

QY 770 IleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMet 789
DB 779 ATAACAGATTTTACTTGAAGACACTCCAGACA-TGCCGGATATGTGCGAGGCTTGCATG 837

QY 790 TyrGluCysArgGluCysTyrAspAspPro-AspIleSerAlaGlyLysIleLysGlnPh 809
DB 838 TATGANTGTANAGAANTGCTACGACGATCCGGGACATCTCAGCTGGAATAAATCAAGCAGTT 897

QY 809 eCysLysThrCysAsnThrGlnValHisLeuHis-ProLysArgLeuAsnHisLysTyrA 829
DB 898 TTGTAACCTGCAACACTCCAGTCCACTTCTTCCCGAGAGCTGAATCTTAATTTT 957

QY 829 snProVal-SerLeuProLysAspLeuProAspTrp-AspTrpArg 843
DB 958 ACCCGTGTGCTCCCTTCCAAAGACTACCCGCACTGGGAGCTGGAAG 1003

RESULT 8
BO437960 1039 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6411653 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5530295
DEFINITION 5', mRNA sequence.
ACCESSION BM457960
VERSION BM457960.1 GI:18507000
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```


DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12210 row: b column: 24
 High quality sequence stop: 702.
 Location/Qualifiers
 source
 1..1039
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5530295"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 3,44e-151 Length: 1039
 Score: 1570.00 Matches: 312
 Percent Similarity: 91.35% Conservative: 5
 Best Local Similarity: 89.91% Mismatches: 16
 Query Match: 31.19% Indels: 14
 DB: 4 Gaps: 4

US-09-671-687A-3 (1-949) x BM457960 (1-1039)

QY 547 GlnileGluAArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGlu 566
 |||||
 Db 14 CAGATTGAGCGGTGAACCTTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAA 73
 |||||
 QY 567 AsnThrProProlyMetGluGlyGluGlyLeuGluMetileGlyLysLysLysGly 586
 |||||
 Db 74 AATACTCCACCAAAATGAAAAAGAGGCTCGAGATATGATGTTGGGAAGAAAGG 133
 |||||
 QY 587 IleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPhe 606
 |||||
 Db 134 ATCCAGGGTCATTACATCTTGTTACTTAGACTCAACCTTATTCGCTTATTTGCTTTT 193
 |||||
 QY 607 SerSerValLeuAspThrValLeuLeuAArgProLysGluLysAsnAspValGluTyrTyr 626
 |||||
 Db 194 AGTCTGTTCTGACACTGTTTACTTAGACCCAAAGAAAGCAACATGTAGATATATAT 253
 |||||
 QY 627 SerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyr 646
 |||||
 Db 254 AGTGAACCCAGAGCTACTGAGGACAGAAATTTGTAATCCTCTGAGATATATGATAT 313
 |||||
 QY 647 ValCysAlaThrLysIleMetLysLeuAArgLysIleLeuGluLysValGluAlaAlaSer 666
 |||||
 Db 314 GTGTGTGCCACAAAATTTAGAACTGAGGAAATATCTTGAAGGTGGAGGCTGATCA 373
 |||||
 QY 667 GlyPheThrSerGluGluLysAspProGluGluPheLeuAsnileLeuPheHisIle 686
 |||||
 Db 374 GGATTACCTCTGAAGAAAAGATCTGAGGAATCTTGAATATCTGTTTCATCATATT 433
 |||||
 QY 687 LeuArgValGluProLeuLeuLysIleAArgSerAlaGlyGlnLysValGlnAspCysTyr 706
 |||||
 Db 434 TTAAGGGTGAACCTTTGCTTAAATAAAGATCAGCAGGTCAAAAGGTACAAGATTGTTAC 493
 |||||
 QY 707 PheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeu 726
 |||||
 Db 494 TTCTATCAAAATTTTATGAAAAAATGAGAAAGTTGGGCTTCCCAATTCAGCAGTTG 553
 |||||
 QY 727 LeuGluTyrSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIle 746
 |||||
 Db 554 TTAGAATGCTCTTTTATCAACAGTACCTGGAATTTGCAGAGCACCATCATGCTCATTT 613
 |||||
 QY 747 IleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeu 766
 |||||
 Db 614 ATTCAGATGCTCGATTGGAAAAAGACTTAAACTATTTAAAAAATTTTCTCTCTG 773

QY 767 GluIleuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGly 786
 |||||
 Db 674 GAATTAATAATAACAGATTTTACTTTGAAGACACTCCACAGACATGCGGATATGTGAGGG 733
 |||||
 QY 787 LeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIle 806
 |||||
 Db 734 CTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793
 |||||
 QY 807 LysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHis 826
 |||||
 Db 794 AAGCAGTTTTGTAAACCTGCAACACACTCAAGTCCACCTTTCATCCGAGAGGCTGAATCAT 853
 |||||
 QY 827 LysTyrAsn-ProValSerLeuProLysAsp-LeuProAspTyr-AspTyr-ArgHisGly 845
 |||||
 Db 854 AATATAAACCCAGTGTGCTTCCCAAAGACTTTTACCCGACTGGGACTGGNAGACACGG 913
 |||||
 QY 845 YCysIlePro--CysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHis 864
 |||||
 Db 914 CTTGCATCCCTTGGCCAGAAATATG-----GGAGTTATTTTGTGCTCTCCGCCA 964
 |||||
 QY 864 sTyrValAlaPheValLysTyrGlyLysAspSerAlaTyrLeuPhePheAspSerMe 884
 |||||
 Db 965 TAGAAAACAA--GCCACCTATGCT-----TGGCTTTTGTGAGATAT 1006
 |||||
 QY 884 tAlaAspArgAsp 888
 |||||
 Db 1007 GTGGAAAGGAAC 1019

RESULT 9
 LOCUS AK042764
 DEFINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:A730022C13 product:CYLINDROMATOSIS (TURBAN
 TUMOR SYNDROME) homolog [Homo sapiens], full insert sequence.

ACCESSION AK042764
 VERSION AK042764.1 GI:26335312
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 3 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 4 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 5 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection

JOURNAL REFERENCE	Nature 409, 685-690 (2001)	QY	1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	DB	190 ATGAGTTTCAGGCTGTGGAGCCCAAGAAAGTTTACTTCACCTTACTCGGAAGAACGGATT 249
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	QY	21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
JOURNAL REFERENCE	Nature 420, 563-573 (2002)	DB	250 TTTTATCTGCTTCTTCAGNATGCGTGTAAACAGCAACAACTCAGAGCTGCTGAAA 309
REFERENCE	6 (bases 1 to 1478)	QY	41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tsegawa, A., Takanashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	DB	310 GTACCCAAAGGGAGTAGGACAGTACATCCAAAGACGTTCTGTGGGGCATTCAGAGTT 369
TITLE	Direct Submission	QY	61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	DB	370 CCTTCACAAAAGGCAAGAAATATCAGATTGGATTAAAAATCTTGGGACCAACCGCATGCA 429
COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	QY	81 ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeuLeu 99
FEATURES	source	DB	430 GTTCTGTTTGTGATGAAAAGGATGTTGTAGAATAAATAAATAATTCACAGAGTTACTG 489
	Location/Qualifiers	QY	100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
	1. .1478	DB	490 TTGGCAATTACCAACTGTGAGGAGAGGCTCAGCCTATTTAGAAACACACTCCGACTAAGT 549
	/organism="Mus musculus"	QY	120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
	/mol_type="mRNA"	DB	550 AAAGGCTCCAGGTAGACGTGGGCGAGTCTGTGAAAGTACAGCTGCGATCTGGGGAAGAG 609
	/strain="C57BL/6J"	QY	140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
	/db_xref="FANTOM DB:A730022C13"	DB	610 AAATTTCCAGGAGTTGTACCTTCAGAGGACCTTTATTAGCGGAGAGACCGTGTGGGG 669
	/db_xref="taxon:10090"	QY	160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
	/clone="A730022C13"	DB	670 ATTTTCTTTGAGTAGAATTTATTTGGAAGAAGTCTGTGTCAGGTTTCACGGATGGGTA 729
	/tissue_type="cerebellum"	QY	180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
	/clone_lib="RIKEN full-length enriched mouse cDNA library"	DB	730 TATCAAGGGAAGCAGCTTTTCCAGTGTGATGAGGACTGTGGCGTGTGTTGTTGTCATTTGAC 789
	/dev_stage="7 days neonate"	QY	199 LysLeuLeuLeuLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
	190. .1146	DB	790 AAATCTAGAATTTATAGAAATGATGACAAATGGATTGGAAAGTGAATTTTTCAGGCGCCAGGA 849
CDS	/notes="unnamed protein product; CYLINDROMATOSIS (TURBAN TUMOR SYNDROME) homolog (Homo sapiens) (SPTR Q96EH0, evidence: PASTY, 93.8%ID, 30.5%length, match=873) putative"	QY	219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
	/codon_start=1	DB	850 GATACATATGACAGTGTGGACCTCCCTTTGGAATAAATCCAGAGTTCTTTTGAAGTT 909
	/protein_id="BAC311357.1"	QY	239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
	/db_xref="GI:263335313"	DB	910 GGAGAAAGTACAGAACTCTGGAACAGTAAATATCTGTGATGTTTACCAGGAAAAGAGAGT 969
	/translations="MSSGLMSQKSTPYWEERIFYLLQECSTDKOTQKLKVPKG SIGQYLDQSVGHSRVPSTGKQKQJLKEPQHAVLFVDEKDVVEINEKFTLELLA IINCERLLFRNRLRLSKGLQVDSVPKQVRSRSEKPPGVVRFRGPLLARTYSG IFFGVLLBEGRGQGTGQVYQKQLFQCDQGVFALDKLELIDDDNGLSDSDFAG PGDTMQVPEPPPEINRSVLKVGESTSGTVFCDVLPKGSLEGLYFVGVDMDNPIGNW DGRFDGVLQCSFASVESTILLHINDIIPGTSKNILQDLKQK"	QY	259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
		DB	970 CTAGGATATTTGTTGTTGTTGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
		QY	279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
		DB	1030 GATGGAGTACAGCTCTCTAGTTTTCAGAGTGTGAAAGTACAAATTCCTCGACATCAAT 1089
		QY	298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
		DB	1090 GACATCATCCAGGGACT-----TCC 1110
ORIGIN		QY	318 ArgGlyValGlyAsp-----LysGlySerSerSerHisAsnLysProLysAlaTh 334
Alignment Scores:		DB	1111 AAGAACAATCTTGTACAGCAACTCAAAGGGAGTAGCTCACAC-----CCTTAGCATTG 1164
Pred. No.:	1.75e-139	QY	334 rGlySerThrSerAspProGlyAsnArgArgSerGluLeuPheTyrThrLeuAsn 352
Score:	1459.50	DB	1165 GGAGGACCTCATCTACCCA-----CACAGAGGACCACCTATCTTTACAATAAGC 1213
Percent Similarity:	86.63%		
Best Local Similarity:	83.57%		
Query Match:	28.99%		
DB:	3		
US-09-671-687A-3 (1-949) x AK042764 (1-1478)			
RESULT	10		

BI666276	603290787F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5325140 5', mRNA sequence.	938 bp	linear	EST 12-SEP-2001
LOCUS	BI666276			
DEFINITION	mus musculus (house mouse)			
ACCESSION	BI666276.1	GI:15580509		
VERSION	EST.			
KEYWORDS	Mus musculus			
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
ORGANISM	NIH-MGC http://mgc.nci.nih.gov/.			
REFERENCE	National Institutes of Health, Mammalian Gene Collection (MGC)			
AUTHORS	Unpublished (1999)			
JOURNAL	Contact: Robert Strausberg, Ph.D.			
COMMENT	Email: cgabbs@mail.nih.gov			
	Tissue Procurement: Jeffrey Green M.D.			
	cDNA Library Preparation: Life Technologies, Inc.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLAM1824 row: n column: 21			
	High quality sequence stop: 776.			
FEATURES	Location/Qualifiers			
source	1. .938			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/strain="FVB/N"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:5325140"			
	/sex="female, virgin"			
	/tissue type="infiltrating ductal carcinoma"			
	/dev stage="5 months"			
	/lab_host="DH108"			
	/clone_lib="NCI_CGAP Mam6"			
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"			
ORIGIN				
Alignment Scores:				
Pred. No.:	2,79e-137	Length:	938	
Score:	1435.00	Matches:	277	
Percent Similarity:	93.16%	Conservative:	9	
Best Local Similarity:	90.23%	Mismatches:	18	
Query Match:	28.51%	Indels:	5	
DB:	4	Gaps:	1	
US-09-671-687A-3 (1-949) x BI666276 (1-938)				
Qy	637 lleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	556		
Db	2 ATAGTCAATCTCTGAGAATATATGATATGTGTGCCACAAAGATTATGAACTGAGG	61		
Qy	657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysAspProGlu	676		
Db	62 AAAATCTTGAAGAGTTGAGGCTGCATCAGGTTTACCTCTGAAGAAAGATCTCGAA	121		
Qy	677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg	696		
Db	122 GAATCTTAATATCTCTGTTTCATGATATTTTAAGGTTGAACCATTTTAAANAAGA	181		
Qy	697 SerAlaGlyGlnLysValGlnAsp-CysTyrPheTyrGlnIlePheMetGluLysAsnG	716		
Db	182 TCAGCAGGTCAAAAGTTCAAGACGGTAACTTCTCAAAATTTTATCGAAAAAATGA	241		
Qy	716 uLysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLe	736		
Db	242 GAAAGTTGGAGTACCCACAATTGAGAGTTATTAGAAATGCTCTTTTATCAACGACCA	301		
Qy	736 uLysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPh	756		
Db	302 GAAATTTGACAGAGCACCATCATCTTGATTATCCAGATGCTCGGTTNGAAAGACTT	361		
Qy	756 eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs	776		
Db	362 TAAACTATTAAAAAATTTTCTCTCCCTGGAATTAATATACAGATTACTTTGAAGA	421		
Qy	776 pThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTy	796		
Db	422 CACTCCAGGACAGTCCGCATCTGTGGAGGACTTGCAGATGACGAGTGCAGGAGTGCTA	481		
Qy	796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl	816		
Db	482 TGACGATCCGACATCTCAGCTGGGAAGATCAAGCAGTTCTGTAAAGACTTCGACGAC	541		
Qy	816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs	836		
Db	542 GGTTCACCTTCATCCAGGAGTTGAATCATCTTATCATCCAGTATCATTCCCAAGA	601		
Qy	836 pLeuProAspTrpAspTrpArgHisGlyCysIlePheProCysGlnAsnMetGluLeuPheAl	856		
Db	602 CTTGCTGACTGGGACTGGAGACACGCGCTGCATCCCTGTGAGAAGATGGAGTTATTTC	661		
Qy	856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyr-GlyLysAspAspS	876		
Db	662 TGTTCCTGTCATGAAACGAGCCACTATGTAGCTTTTGTGAAGTATGGGGAAGGATG	721		
Qy	876 erAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI	896		
Db	722 CTGCTGGCTTTTCTTTGACAGATGGCGATCGAGATGGTGGTCAAGATGGTTCAAGA	781		
Qy	896 leProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspL	916		
Db	782 TTCCACAGTGCAGCTGCGCCAGAGTGGGAGAGTACTGAAGATGCTCTCGGGAAGAAC	841		
Qy	916 eutHisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlar	936		
Db	842 TGACCTCTGGGACCC-AAAAGAATCCAGGCGC---GGCCCGCAACA-CTAGGCGAGG	896		
Qy	936 yrMetCysMetTyrGln	941		
Db	897 ACATGTGCATGTACCAG	913		
RESULT 11				
AK048183	2218 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone: C13009D01 product: CYLINDROMATOSIS (TURBAN TUMOR SYNDROME) homolog (Homo sapiens), full insert sequence.			
DEFINITION				
ACCESSION	AK048183			
VERSION	AK048183.1	GI:26092705		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			

AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE PUBLISHED	Genome Res. 10 (11), 1757-1771 (2000)
20530913	
11076861	
4	
REFERENCE AUTHORS	THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL REFERENCE AUTHORS	Nature 409, 685-690 (2001)
5	
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL REFERENCE AUTHORS	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
6	
Nature 420, 563-573 (2002)	
(bases 1 to 2218)	
ADACHI,J., AIZAWA,K., AKIMURA,T., ARAKAWA,T., BONO,H., CARNINCI,P., FUKUDA,S., FURUNO,M., HANAGUCHI,T., HARA,A., HASHIZUME,M., HAYASHIDA,K., HAYATSU,N., HIRAMOTO,K., HIRAOKA,T., HIROZANE,T., HORI,F., IMOTANI,K., ISHII,Y., ITOH,M., KAGAWA,I., KASUKAWA,T., KATO,H., KAWAI,J., KOJIMA,Y., KONDO,S., KONNO,H., KOUNDA,M., KOYA,S., KURIHARA,C., MATSUYAMA,T., MIYAZAKI,A., MURATA,M., NAKAMURA,M., NISHI,K., NOMURA,K., NUMAZAKI,R., OHNO,M., OHEATO,N., OKAZAKI,Y., SAITO,R., SAITOH,H., SAKAI,C., SAKAI,K., SAKAZUME,N., SANO,H., SAKAI,D., SHIBATA,K., SHINAGAWA,A., SHIRAKI,T., SOGABE,Y., TAGAMI,M., TAGAWA,A., TAKAHASHI,F., TAKAKU-AKAHIRA,S., TAKEDA,Y., TANAKA,T., TOMARU,A., TOYA,T., YASUNISHI,A., MURAMATSU,M. and HAYASHIZAKI,Y.	
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC);, RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp).
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/.
FEATURES	Location/Qualifiers
source	1. 2218 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:C130039D01" /db_xref="taxon:10090" /clone="C130039D01" /tissue_type="head" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="16 days embryo"
misc_feature	1. 2218 /note="CYLINDROMATOSIS (TURBAN TUMOR SYNDROME) homolog (Homo sapiens) (SPTRLQ96EH0, evidence: FASTV, 93.8%ID, 30.5%length, match=873)"
ORIGIN	
Alignment Scores:	
Pred. No.:	1.03e-135 Length: 2218
Score:	1426.00 Matches: 273
Percent Similarity:	96.56% Conservative: 8

KEYWORDS	EST.	Source	Location/Qualifiers	Length	Matches	Score	Percent Similarity	Best Local Similarity	Query Match	DB
ORGANISM	Homo sapiens (human)			3,17e-135	1133					
REFERENCE	1 (bases 1 to 1133)									
AUTHORS	NIH-MGC									
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)									
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM12112 row: h column: 16 High quality sequence stop: 747.									
FEATURES	1..1133									
source	/organism="Homo sapiens"									
	/mol_type="mRNA"									
	/db_xref="taxon:9606"									
	/clone="IMAGE:5491647"									
	/tissue_type="retinoblastoma"									
	/lab_host="DH10B (phage-resistant)"									
	/clone_lib="NIH MGC 67"									
	/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."									
ORIGIN										
Alignment Scores:										
Pred. No.:	3,17e-135			Length:	1133					
Score:	1416.50			Matches:	307					
Percent Similarity:	86.76%			Conservative:	14					
Best Local Similarity:	82.97%			Mismatches:	23					
Query Match:	28.14%			Indels:	26					
DB:	4			Gaps:	9					
US-09-671-687A-3 (1-949) x BM480127 (1-1133)										
Qy	111	LeuPheLysAsnArgLeuSerLysGlyLeuGlnIleAspValGlyCysProVal	130							
Db	13	CTGTTTAAACAGAAACAGACTAGTAAAGCCCTCCAAATAGACGTGGGCTGCTCTGTG	72							
Qy	131	LysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgGlyPro	150							
Db	73	AAAGTACACTGATCTCGGGAGAGAAATTTCTCGAGTTGTACGTTAGAGGACCC	132							
Qy	151	LeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuGluGly	170							
Db	133	CTGTAGCAGAGAGGACACTCCGGAATATTTCTTTGGAGTTGAATTCGCGAAGGT	192							
Qy	171	ArgGlyClnGlyPheThrAspGlyValTyrglnGlyLysGlnLeuPheGlnCysAspGlu	190							
Db	193	CGTGGTCAAGAGTTTCACTACCGGGGTGTACCAAGGAGAAACACGCTTTTTCAGTGTGATGA	252							
Qy	191	AspCysGlyGlyPheValAlaLeuAspLysLeuGluLeuLeuAspAspThrAla	209							
Db	253	GATTGTGGCGTGTGTTGTCATTGGACAGCTAGAACTCATAGAGATGATGACACTGCA	312							
Qy	210	LeuGluSerAspTyraAlaGlyProGlyAspThrMetGlnValGluLeuProLeuGlu	229							
Db	313	TTGGAAAGTGATTACCCAGGCTCTCGGGACACAATGCAGGTGCAACTTCTCTTTGGAA	372							
Qy	230	IleAsnSerArgValSerLeuLysGlyGlyGluThrIleGluSerGlyThrValIlePhe	249							
Db	373	ATAAATCCAGAGTTTCTTTGAAGGTTGGAGAAACAAATAGAAATCTCGAACAGTTATATTC	432							
Qy	250	CysAspValLeuProGlyLysGluSerLeuGlyTyrrPheValGlyValAspMetAsp	269							
Db	433	TGTGATGTTTGGCAGGAAAGAAAGCTTAGGATATTTTGGTGTGGACATGACATAC	499							
Qy	270	ProIleGlyAsnTrpAspGlyArgPheAspGlyValLLeuCysSerPheAlaCysVal	288							
Db	493	CTATTGGCAACTGGGATGGAAGATTGATGAGTGCAGCTTTGTAGTTTTCGGTGTGT	552							
Qy	289	GluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValThrGlnGluArg	308							
Db	553	GAAAGTACAATTCATTGTCACATCATGATATCATCCAGAGAGTGTGACGAGGAAAGG	612							
Qy	309	ArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySerSerHis	328							
Db	613	AGGCCTCCCAAACTTCGCTTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCAT	672							
Qy	329	AsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArgLLeuSerGluLeu	347							
Db	673	AATAAACCAAGGCTACAGATCTACTCAGACCCCTGGAAATAGAAACAGATCTGAATTA	732							
Qy	348	PheTyrrThrLeuAsnGlySerValAspSerGlnProGlnSerLysSerLysAsnThr	367							
Db	733	TTTATACCTTNAATGGCTCTCTGTTGACTCACAACCAATCC								

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).

Seq primer: M13 REVERSE.

FEATURES

source

Location/Qualifiers
1. 812
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-FJ0-cga-C-11-0-UI"
/tissue_type="embryo"
/dev_stage="embryo"
/lab_host="PH10B (Life Technologies) (T1 phase resistant)"
/clone_lib="UI-R-FJ0"
/note="Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
UI-R-FJ0 is a cDNA library containing the following
tissue(s): rat embryo. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pTVT3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CATCTCTACT. This library
was created for the University of Iowa Program for Rat
Gene Discovery and Mapping (Val Sheffield, Bento Soares
and Tom Casavant)"

ORIGIN

Alignment Scores:

Pred. No.: 4,19e-132 Length: 812
Score: 1384.00 Matches: 256
Percent Similarity: 98.48% Conservative: 4
Best Local Similarity: 96.97% Mismatches: 4
Query Match: 27.49% Indels: 0
DB: 6 Gaps: 0

US-09-671-687A-3 (1-949) x CA512526 (1-812)

627 SerGluThrGlnGluLeuLeuArgThrGluLeuValAsnProLeuArgIleTyrGlyTyr 646
:::|||||
20 ACGGAGACTCAAGAGCTACTGAGGACAGAGATAGTCAATCTCTGAGAATATATGGATAT 79
647 ValCysAlaThrIlystleMetIysLeuArgIlyleleuGluIysValGluAlaAsr 666
:::|||||
80 GTGTGTGCCACAAAGATTATGAAGCTGAGGAAATACTTGAAAAGTTGAGGCTGCATCA 139
667 GlyPheThrSerGluGluIysAspProGluGluPheLeuAsnIleLeuPheHisIle 696
:::|||||
140 GATTTTACTCTGAGGAAAAGATCTCTGAGAATTTCTAAACATCTCTTTTCATGATATT 199
687 LeuArgValGluProLeuLeuIlystleIleArgSerAlaGlyGlnIlyValGlnAspCysTyr 706
:::|||||
200 TTAAGGTTTGAACCATTTGTTAAATAAAGTCAAGCTCAGAGGTCAGAAAGTTCAAGACTGTAA 259
707 PheTyrGlnIlePheMetGluIysAsnGluIysValGlyValProThrIleGlnGlnLeu 726
260 TTCTATCAAAATTTTATGGAATAAATGAGAAAGTCGGAGTACCCCAATCCAGCATTTA 319
727 LeuGluTyrSerPheIleAsnSerAsnLeuIysPheAlaGluAlaProSerCysLeuIle 746
320 TTAGAATGGTCTTTTATCAACACACCTGAAATTTTCAGAGGACCAATCATCTGCTGATT 379
747 IleGlnMetProArgPheGlyIysAspPheIysLeuPheIysIlystlePheProSerLeu 766
:::|||||

Db 380 ATCCAGATGCCTCGGTTTGGGAAGACTTTAAACTATTAAAAAATTTTCTTCTCCTG 439
Qy 767 GluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGly 786
Db 440 GAATTAATATAACAGATTACTTTGAAGACACTCCAGGCGAGTGGCGCATCTGTGGAGGA 499
Qy 787 LeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyIle 806
Db 500 CTCGCCATGTATGAGTGTAGAGAGTCTATGATGACCCGACATCTCGGAGGGAAGATC 559
Qy 807 LysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHis 826
Db 560 AGCAGTTCTGTAAAGACCTGCAGCACTCAGTTCACTTCAATCCAGAAAGACTGAATCAC 619
Qy 827 LysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCys 846
Db 620 ACTTACCATCCAGTATCACTTCCAAAGACTTCCCGAGCTGGGACTGGAGACACGGCTGC 679
Qy 847 IleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrVal 866
Db 680 ATCCCGTGTGAGAAGATGGAGTTATTGTCTGCTCTGCATAGAAACACGACCATATGTT 739
Qy 867 AlaPheValIlystlyrGlyIysAspAspSerAlaTyrLeuPheAspSerMetAlaAsp 886
Db 740 GCTTTTGTGAAGTACGGAAGGATGACTCTGCCTGCTCTTTTGACAGATGGCTGAT 799
Qy 887 ArgAspGlyGly 890
Db 800 CGAGATGGTGGT 811

RESULT 14

BC024596

LOCUS

DEFINITION

Mus musculus cyclinomatosis (turban tumor syndrome), mRNA (cdna

clone IMAGE:4022147), with apparent retained intron.

ACCESSION

BC024596

VERSION

BC024596.1

KEYWORDS

HTC.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 1892)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F.,

Diatchenko, U., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.D.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,

Schnerk, J., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

2 (bases 1 to 1892)

AUTHORS

Strausberg, R.

Direct Submission

Submitted (01-MAR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 41 Row: e Column: 18
 This clone has the following problem: retained intron.

FEATURES
source

1..1892
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4022147"
 /tissue_type="Mammary tumor metastasized to lung."
 MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer."
 /clone_lib="NCI CGAP_Lu30"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:
 Pred. No.: 2,65e-129 Length: 1892
 Score: 1363.00 Matches: 246
 Percent Similarity: 98.81% Conservative: 3
 Best Local Similarity: 97.62% Mismatches: 3
 Query Match: 27,08% Indels: 0
 DB: 3 Gaps: 0

US-09-671-687A-3 (1-949) x BC024596 (1-1892)

QY 698 AlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLys 717
 DB 1 GCAGGTCAAAAAGTTCACAGCTGTACTTCTATCAAAATTTTATGGAAAAAATGAGAAA 60
 QY 718 ValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeuLys 737
 DB 61 GTTGGAGTACCCACAAATTCAGCAGTTATTAGAAATGGTCTTTATCAACAGCAACCTGAAA 120
 QY 738 PheAlaGluAlaProSerCysIleuIleGlnMetProArgPheGlyLysAspPheLys 757
 DB 121 TTTCAGAGGCACCATCATGCTGTGATTATCCAGATGCCCTCGGTTTGGAAAAAGACTTTAAA 180
 QY 758 LeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThr 777
 DB 181 CTATTATAAAAAATTTTCTCCCTCGGAATTAATATACAGATTACTTGAAGACACT 240
 QY 778 ProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAsp 797
 DB 241 CCCAGGCAGTCGGCATCTGTGGAGGACTTGGATGTACGAGTGCAGGAGGTCTATGAC 300
 QY 798 AspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnVal 817
 DB 301 GATCCGGACATCTCAGCTGGGAAGATCAAGCAGTTCTGTAAAGACCTGCAGCACTCAGGTT 360
 QY 818 HisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeu 837
 DB 361 CACCTTCATCCCGAGGAGGTGAATCTTCTATCATCCAGTATCACTTCCCAAGACTTG 420
 QY 838 ProAspTyrAspTrpArgHisGlyCysIleGlyCysGlnAsnMetGluLeuPheAlaVal 857

Db 421 CTTGACTGGAGCTGGAGACACCGCTGCATCCCTGTGAGAGATGGAGTTATTTCGTGT 480
 QY 858 LeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAla 877
 Db 481 CTCCTGATAGAAACGAGCCACTATGTTGCTTTTGTGAAGTATGGAGGATGACTCTGCC 540
 QY 878 TrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIlePro 897
 Db 541 TGGCTTTTCTTTTACACAGCATGGCGATCGAGATGGTGTGAGATGGCTTCAACATTTCA 600
 QY 898 GlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHis 917
 Db 601 CAAGTGACGCCCTGCCAGAGTGGGAGAGTACTTTGAAGATGCTCTGGAGGAGCTGCAC 660
 QY 918 SerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMet 937
 Db 661 TCTTTGATTCAGAGAGGATTCAGGCTGTGGCGCAGACTTCTTTGGATGCATACATG 720
 QY 938 CysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 721 TGCATGTACCAGAGTCCACCATGAGCCTGTACAAA 756
 RESULT 15
 BG867631
 LOCUS
 DEFINITION
 mRNA sequence.
 ACCESSION
 BG867631
 VERSION
 BG867631.1 GI:14218171
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 837)
 AUTHORS
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10818 row: 1 column: 03
 High quality sequence stop: 756.
 FEATURES
 Location/Qualifiers
 1..837
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4913426"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP_SG2"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1: NCI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:
 Pred. No.: 1,59e-127 Length: 837
 Score: 1340.00 Matches: 254
 Percent Similarity: 95.26% Conservative: 7
 Best Local Similarity: 92.70% Mismatches: 12
 Query Match: 26.62% Indels: 2
 DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x BG867631 (1-837)

```
Qy 611 AspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGln 630
Db 2 GACACTGTGTTACTTAGACCCAAAGAGAAGATGATATAGAGTATTACAGTGAGACTCAG 61
Qy 631 GluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThr 650
Db 62 GAGCTACTGAGGACAGAGATAGTCAATCCTCGAGATATATGGATATGTGTGCCACA 121
Qy 651 LysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSer 670
Db 122 AAGATTATGAACCTGAGGAATACTTGAAAAGTTGAGGCTGCATCAGGTTTACCTCT 181
Qy 671 GluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGlu 690
Db 182 GAAGAAAAGATCCTGAAGATTCTTAATATCCTGTTTCATGATATTTTAAGGGTTGAA 241
Qy 691 ProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIle 710
Db 242 CCATTGTTAAAAATAAGATCAGCAGGTCAAAAAGTTCAAGACTGTAACTTCTATCAAAAT 301
Qy 711 PheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuTrpSer 730
Db 302 TTTATGGAAAAAATGAGAAAGTTGGAGTACCCCAATTCAGCAGTTATTAGAATGGTCT 361
Qy 731 PheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetPro 750
Db 362 TTTATCAACAGCAACTGAATTTGCAGAGGACCACATCATGTTGATATCCAGATGCCT 421
Qy 751 ArgPheGlyLysAspPheLysLeuPheLysIlePheProSerLeuGluLeuAsnIle 770
Db 422 CGGTTCCGAAAAAGACTTTAAACTATTAAAAAATT-TTTCCTTCCCTGGAAATTAATAATA 480
Qy 771 ThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyr 790
Db 481 ACAGATTTTACTTGAAGACACTCCAGGCAGTCCGCATCTGTGGAGGACTTGGCATGTAC 540
Qy 791 GluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCys 810
Db 541 GAGTGCAGGGAGTGCTATGCGATCCGGACATCTCAGCTGGGAAGATCAAGCAGTTCTGT 600
Qy 811 LysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnPro 830
Db 601 AAGACTGCAGCACTCAGGTTCACTTCATCCAGGAGGTGAATCATCTTATCATCCA 660
Qy 831 ValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGln 850
Db 661 GTATCACTTCCCAAAGACTTGCTGACTGGGACTGGAGACACGGCTGCATCCCTGTCA 720
Qy 851 AsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLys 870
Db 721 AGATGGAGATTATTGCTGTTCTGATAGAAACGAGGACCTATGTCGCTTTGTGAAG 780
Qy 871 TyrGlyLys-AspAspSerAlaTrpLeuPhePheAspSer 883
Db 781 TATGGGAACGGATGACTCTGGCTGGCTTTCTTTCGACACG 820
```

Search completed: April 18, 2005, 20:07:09
Job time : 6231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:18:59 ; Search time 22 Seconds
(without alignments)
4150.441 Million cell updates/sec

Title: US-09-671-687A-3
Perfect score: 5034
Sequence: 1 MSSGLWSQEKVTSPTWEERI.....RLLCDAYMCVQSPMSLYK 949

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	613.5	12.2	1021	2 F88568	protein F40F12.5 [
2	590.5	11.7	727	2 S42834	F40F12.5 protein -
3	151	3.0	1392	2 A43336	microtubule-vesicl
4	151	3.0	1427	2 S22695	restin - human
5	149.5	3.0	1046	2 T42734	cytoplasmic linker
6	149	3.0	1046	2 T42734	cytoplasmic linker
7	135.5	2.7	2273	2 T09083	hemagglutinin/hemo
8	130	2.6	717	2 S38177	SSV7 protein homol
9	125.5	2.5	1921	2 T13827	kinesin-73 - fruit
10	124.5	2.5	1146	2 F83853	5-methyltetrahydro
11	124	2.5	3624	2 AD0835	large repetitive p
12	122.5	2.4	1953	2 S63244	BN1L protein - yea
13	122	2.4	1791	2 T02345	hypothetical prote
14	121	2.4	1498	2 E86302	hypothetical prote
15	121	2.4	1641	2 T10955	early nodulin bind
16	120	2.4	499	2 S56265	hypothetical prote
17	119	2.4	662	2 T18233	probable transcrip
18	118.5	2.4	1466	2 T17138	CL1AA protein - ra
19	118.5	2.4	1467	2 T18411	latrophilin-1, bra
20	118.5	2.4	1471	2 T17149	CL1BA protein - ra
21	118.5	2.4	1472	2 T18413	latrophilin-1, bra
22	118.5	2.4	1510	2 T17145	CL1AB protein - ra
23	118.5	2.4	1515	2 T17156	CL1BB protein - ra
24	116.5	2.3	825	2 T47164	hypothetical prote
25	116.5	2.3	2484	2 T26216	hypothetical prote
26	116.5	2.3	2607	2 T25509	hypothetical prote
27	115.5	2.3	514	2 T25509	hypothetical prote
28	115.5	2.3	670	2 T38446	microtubule-associ
29	115.5	2.3	673	2 S35335	transcription fact

30	115.5	2.3	1341	2 T17285	hypothetical prote
31	115	2.3	471	2 S57591	hypothetical prote
32	114.5	2.3	752	2 D96699	hypothetical prote
33	114.5	2.3	1161	2 B70387	DNA polymerase III
34	114.5	2.3	1357	2 S61187	probable membrane
35	114	2.3	438	2 T39772	ubiquitin carboxyl
36	114	2.3	739	2 T15215	hypothetical prote
37	113.5	2.3	974	2 D89057	protein K09H1.1 [
38	113.5	2.3	1335	2 T18289	racGAP protein - s
39	113.5	2.3	2700	2 D88450	protein F21H1.2 [
40	113	2.2	753	2 A96747	probable RNA-bind
41	113	2.2	1226	1 S65593	adenosine deaminas
42	112.5	2.2	642	2 T10861	phaseolin G-box bi
43	112.5	2.2	805	2 S50277	ubiquitin-specific
44	112.5	2.2	1116	2 I54378	gene X104 protein
45	112	2.2	3685	1 A27605	dystrophin, muscle

ALIGNMENTS

RESULT 1

F88568
protein F40F12.5 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: F88568
R/anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998

~~October~~

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: F88568
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1021 <STO>
A/Cross-references: GB:chr_III; PIDN:CAA82939.1; PID:g3877002; GSPDB:GN00021; CESP:F40F12.5
C/Genetics:
A/Gene: F40F12.5
A/Map position: 3

Query Match 12.2%; Score 613.5; DB 2; Length 1021;
Best Local Similarity 23.7%; Pred. No. 4.6e-34;
Matches 256; Conservative 154; Mismatches 315; Indels 357; Gaps 46;

Qy	103	TNCEERPSLF-----KVRNRLSKGLQIDVGCVKVQLRSGEEKPFGVVRFRGPLLAERTV	157
Db	43	TSSEKNSNLYTSENTSNRLNLNQNPVGT-----	73
Qy	158	SGIFFGVLEEGRGOGFTDGVYQKGLFOCDEDCGFVALDKLELIEDDDTALE-----	211
Db	74	--LIDGPELINEASGSGFLD-----DQL-----VDVSDYSDRTTKLDRNRSP	115
Qy	212	-----SDYAGPGDTMQVELPPLEI-----	235
Db	116	ELIVALLQKRVQGRFSSNYGREPCVIEHPGTVMREADDKMSKSELKXWFTKSRAS	175
Qy	236	--LKGG-----ETIESGTVIFC-----DVLPGKESLGYFVGVDMDNPIGNWDGRFGCV	281
Db	176	SHLRDGLAMPMEELDC-TPLICAMITRSDVMRINQDQAIHLLAVSVEKRI-----EV	226
Qy	282	LCSEFACVESITLLHINDIIPESV--TQERRPPKLAFMRSRGVGDKG-----	324
Db	227	YQNFENFNFNLNLKIGDSVSEVDETMRRVRPAKVSWTGERPEASGIWYNVDFDGNTSQMP	286
Qy	325	-----SSSHNKP-----KATGSTS-DPCGNRSSELYFTLN-----GSSVDS--	358
Db	287	SSNQSYSSSHDRLNQRQDTWNFMSSGSSVAPSN--SRLYSNPQMMPMKGGVGSALY	344
Qy	359	-----OPQSKSKNTWYIDEVAEPKSLTRIS	385
Db	345	DNRRLVQYSGDEEQYRSAPKAPRRIIPVSRQOPEIQRN-----SRSMKPSSE	393

[illegible]

Db 7 KGSISLIFFCHSGSSVAPSN--SRLYYSNOMHPMKGGVGLVADNRRLLVQVSGDEEQ 64
 QY 359 -----QPOSKSKNTWYIDEVAEDPAKSLTEISTDFDR-SSPPLQPP 398
 Db 65 YRSAPKAPRERIIPVSRQOPELEQRN-----SRMKPSEFDYNTYSTHPRPP 113
 QY 399 PVNSLT--TENRPHSLFSLTKMNPNT---NGSIGHGPLSLSAOSVMBELNTPVQESPP 453
 Db 114 PSSMMNYPMSNTHSLOPSRSKSVQTIQRNQFVRQNPARNERNVESDQLNFRIGDQC-- 170
 QY 454 AMPPGNSHGLE---VGSIAEVKENPPFY-----GVTRWIGOPPLNEVL 494
 Db 171 ----INWNNGAERGIIKIYIGFLKGHKTLYAGSHVVVRHLGAQRGTGVVWLGE----- 219
 QY 495 AGLELEDECAGCTDGTFRGTRYFTCAKALFKVLKSCRPD-----SRFAS----- 540
 Db 220 ----EIDDE-----GDR-----IVRSAI VQLEDVFTAMRRSHDASYSYTPVVG 266
 QY 541 -LQPVSNQIERCNSLAFGGYLSVEVBENTPPK-----MEKE----- 575
 Db 262 VLVP IQSLKHNHSGSSNG--ASVQVOTSTYPTQTYNNISRRRTBDFGMSDGVKEQKCGI 319
 QY 576 ---GLEIMIGKKXGIGHYNSCYLDSTLFCLPARSSVLDTVLLRPEKNDVVEYSETQELL 633
 Db 320 AKDMOQLVGRQGIQGYCNSCYLDATLYAMFVQVTCDFLL--EKSIGKSETAQOQKIL 377
 QY 634 RTEIVNPLRIYGVVCATKMKLKRIL-EKVEAASGFTSEKDPREFNLIPHHILRVRPL 692
 Db 378 AHEIVFPLKXHVVRADHVMKRLKLAELMPHTGLTNEEKDPEIIGFIPSKVFHAEPP 437
 QY 693 LKRSAGQVQDCYFYQIFMEKNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPPRF 752
 Db 438 IKLIGONHAKDSQYLPVIVD-DWLGSAAFSQHLLRHMERSAQVTFAKAPPVLIMQLPRY 496
 QY 753 GKDFKLPKFIPLPSLELNITDLLEDTPRQIRICGLAMVECRECYDDPDISAGKIKQFCKT 812
 Db 497 GQQ-KYFDKILPLLETIDITPFVAGAVPACSKQACSEVYCPTCLTRRVFYSEV-IFCRK 554
 QY 813 CNTQVHLHPKRLNHYN---PVSLPKDLPDWDWRHGGCIPQNMELFAVLCTETSHYAVF 869
 Db 555 CFHITHLLPEIDHKSRDLYPPGKPKKPH-----SHKMVLSAVLCLETSHYAVV 605
 QY 870 KYGKDSAMLFFPDSMADRGQGNFIPOVTPCPVEGYELKMS-----LEDLH----- 917
 Db 606 R--TSSNQWVFFPDSMADRGSLDGFNVVYVRECGNMSDWLSLQGNRLKDADECGQIKV 663
 QY 918 -----SLDSRRITQ-----GCARLLCDAYCMYQ--SPTWS 946
 Db 664 ELKNKINSPDRKLEIAMFGQSSLDPLVGRLLSDSYICFYEDASPTSS 711
 RESULT 3
 A43336
 microtubule-vesicle linker CLIP-170 - human
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
 A;Title: CLIP-170 links endocytic vesicles to microtubules.
 A;Reference number: A43336; MUID:92405160; PMID:1356075
 A;Accession: A43336
 R;Pierre, P.; Scheel, J.; Richard, J.E.; Kreis, T.E.
 Cell 70, 887-900, 1992
 A;Molecule type: mRNA
 A;Residues: 1-1392 >PIE
 A;Cross-references: GB:G1897501; NID:G180621; PIDN:AAA35693.1; PID:G180622

```
Db      85  GQWAGIVLDEPIKNDGSAVGR-YFQCEP-----LKGF-----TRPSKL---TRK 127
QY      320 VDGKSSSHNKPATGSDPGNRRSELFYTLNGSSVDSPQSKSKNTWYIDVADPAK 379
Db      128 V-----QAEDANGLOTTPASRATSPCTSTASVSSPSPSN-----IPKPSQPA- 175
QY      380 SLTEISTDFDRSSPPLQPPVNSLTTFNPHSLTPKMPNTNGSIGHSPLSLSAQSV 439
Db      176 -----AKEPSATPPISNLT-----KTASESIS 197
QY      440 EELNTAPVOESPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPGLNEVLAGLEL 499
Db      198 NLSEAGSIKK-----GERELKIGDRVVGGTAK--GVRFLGETPAKGEWCGVEL 246
QY      500 EDECACTDGTFRGTRYFTCALKKALFVKLKSRCRPSRFPASLQPVSNQIERCNSLAFGGY 559
Db      247 -DEPLGKNDGAVAGTRYFOQPKYGLFA-----PVHKVTKIGF--- 283
QY      560 LSEVVEENTPPKMEKGL-EIMIGKKKGIQGHYNSCYLDSTLFCLEAFSSVLDTVLLRPK 618
Db      284 -----PSTTPAKAKANAVRRVMTTASLKRSPSASSLS-----MSSVASSVSRP- 330
QY      619 EKNDVEYSETQELLRTIENVPLRIYGYVCATKIMK-----LRKILEKVEAASGP 668
Db      331 -----SRTGLTTSRYARKISGTTALQEALEKQKHIEQLLAERDLERAFAKA- 381
QY      669 TSEKDPPEFLNIL-----FHHIL-----RVEPLLKIRSAQKVQDC 705
Db      382 TSHVGEIEQELALARDHQHVLLEAKMDQLRTWVEAADREKVLELLNQLNEEKKVEDL 441
QY      706 YF 707
Db      442 QF 443

RESULT 4
restin - human
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S22695; S19853
R:Biiber, G.; Delabie, J.; Bruengen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;
EMBO J. 11, 2103-2113, 1992
A:Title: Restin: a novel intermediate filament-associated protein highly expressed in th
A:Reference number: S22695; MUID:92289675; PMID:1600942
A:Accession: S22695
A:Molecule type: mRNA
A:Residues: 1-1427 <BIL>
A:Cross-references: UNIPROT:P30622; EMBL:X64838; NID:g35998; PIDN:CAA46050.1; PID:g35999
C:Keywords: cytoskeleton

Query Match      3.0%; Score 151; DB 2; Length 1427;
Best Local Similarity 20.5%; Pred. No. 0.089;
Matches 99; Conservative 55; Mismatches 171; Indels 157; Gaps 20;

QY      260 GYFVGVDNPTIGNWDGDFGLCSFACVESTILLHINDIIPESVTQERRPPKLAFMERG 319
Db      85  GQWAGIVLDEPIKNDGSAVGR-YFQCEP-----LKGF-----TRPSKL---TRK 127
QY      320 VDGKSSSHNKPATGSDPGNRRSELFYTLNGSSVDSPQSKSKNTWYIDVADPAK 379
Db      128 V-----QAEDANGLOTTPASRATSPCTSTASVSSPSPSN-----IPKPSQPA- 175
QY      380 SLTEISTDFDRSSPPLQPPVNSLTTFNPHSLTPKMPNTNGSIGHSPLSLSAQSV 439
Db      176 -----AKEPSATPPISNLT-----KTASESIS 197
QY      440 EELNTAPVOESPPLAMPNGSHGLEVSLAEVKENPPFYGVIRWIGOPGLNEVLAGLEL 499
Db      198 NLSEAGSIKK-----GERELKIGDRVVGGTAK--GVRFLGETPAKGEWCGVEL 246
QY      500 EDECACTDGTFRGTRYFTCALKKALFVKLKSRCRPSRFPASLQPVSNQIERCNSLAFGGY 559
```

```
Db      247 -DEPLGKNDGAVAGTRYFOQPKYGLFA-----PVHKVTKIGF--- 283
QY      560 LSEVVEENTPPKMEKGL-EIMIGKKKGIQGHYNSCYLDSTLFCLEAFSSVLDTVLLRPK 618
Db      284 -----PSTTPAKAKANAVRRVMTTASLKRSPSASSLS-----MSSVASSVSRP- 330
QY      619 EKNDVEYSETQELLRTIENVPLRIYGYVCATKIMK-----LRKILEKVEAASGP 668
Db      331 -----SRTGLTTSRYARKISGTTALQEALEKQKHIEQLLAERDLERAFAKA- 381
QY      669 TSEKDPPEFLNIL-----FHHIL-----RVEPLLKIRSAQKVQDC 705
Db      382 TSHVGEIEQELALARDHQHVLLEAKMDQLRTWVEAADREKVLELLNQLNEEKKVEDL 441
QY      706 YF 707
Db      442 QF 443

RESULT 5
T42734
cytoplasmic linker protein CLIP-115 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42734
R:de Zeeuw, C.I.; Hoogenraad, C.C.; Goedknegt, E.; Hertzberg, E.; Neubaer, A.; Grosvel
Neuron 19, 1187-1199, 1997
A:Title: CLIP-115, a novel brain specific cytoplasmic linker protein, mediates the local
A:Reference number: 222252; MUID:98087115; PMID:9427243
A:Accession: T42734
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1046 <DEZ>
A:Cross-references: UNIPROT:O55156; EMBL:AJ000485; NID:g2792008; PIDN:CAA04123.1; PID:g2
C:Gene: CLIP-115
C:Function:
A:Description: supposed to operate in the control of brain-specific organelle translocat
A>Note: enriched in the dendritic lamellar body (DLB)

Query Match      3.0%; Score 149.5; DB 2; Length 1046;
Best Local Similarity 20.9%; Pred. No. 0.069;
Matches 115; Conservative 66; Mismatches 189; Indels 179; Gaps 24;

QY      309 RPKLAFMMSGVGDGKSSSHNKPATGSDPGNRRSELFYTLNGSSVDSQ-----PQ 361
Db      8  KPP-----GRGKKHSSPVGRFSI-----GSASSVVASASGSKGSLPHKQASGPS 53
QY      362 SKSKNTWYIDVADPAKSLTEISTDF-----DRSSPPLQPPPVNSLTTFNRF----- 409
Db      54  SAGATT-----TVSEKPGPKAAEVGDDFLGDFVVGGERVWVNGVKFCVVQYLGETQFAPQW 109
QY      410 -----HSLPFSITKMP--NTNGSIGHSPLSLSAQS 437
Db      110 AGVVLDDPVGKNDGAVGGLRYFECALQGIIFTRPSKLTQPAAEAGSGDGHVESLTAQN 169
QY      438 VMEELNTAPVOESPPL-----AMPNGSHGLEVSLAEVKENPPFY----- 478
Db      170 LSLHSGTA-----TPPLTGRVPIRESVINGSVKTNESGNSLSDSGSVKRGDKDLHLGDR 225
QY      479 -----GVIRWIGOPGLNEVLAGLEDEACAGCTDGTFRGTRYFTCALKKALFVKL- 529
Db      226 VLVGGTKTGVRVYVGETDFAKGEWCGVEL--DEPLGKNDGAVAGTRYFOCPKPFGLFAFIH 284
QY      530 -----KSCRPDSRPA-----SLQPVSNQIERCNSLA--FGYILSEVVEENTP 569
Db      285 KVIRIGFPSPSPAKAKTKQAMGVSALTTHSPSSSISSVSSVASSVGRPS----- 336
QY      570 PKMEKGLEIMIGKKKGIQGHYNSCYLDSTLFCLEAFSSVLDTVLLRP--KEKND-VEYY 626
Db      337 -----RSGLLTETSRYA-----RKISGTTALQEALEKQKHIEQL 372
```


Db 1658 SYNEQSGIFAGGQYRIRVNGKTGLVGAUVSD-ADSKNLLKTSIWHKDIQNHASAAA 1716
QY 380 SUTESTDTRSPPIOPPPVNSLTTFNRPFLSLTKMPNTNGSIGHSPLSL-----433
Db 1717 SALGLSGGFSYSPKP-----TSGQYS-----TKAEABIGKIGKPKVSLMRFDQV 1760
QY 434 SAQSVMEELNTAPVQES-----PPLAMPPGNHGLEVG-----466
Db 1761 SAKD--DELNEKYRSRIEKGTFKEANLNQNNAGGLKFLGKQNDIHSNDKYALAKWGLG 1818
QY 467 -SLAEVKNPPFPYGVIRWIGQPPGLNE--VLAGELEDEBACAGCTDGTFRGTRYFTCALX 522
Db 1819 NLLGNAKESQESQITRSV-----ISEGDQWIASAQRKNIAIEKGTSSAHKALAKADR 1873
QY 523 KALFVKLSCRPSRPFASIQVSNQIERNCSLAFGGYLSV-----VEEN 567
Db 1874 EGLKEVELNRDVAK-----EFINETLIGGIADBAVRSQFIAEHLMTFKMDEN 1922
QY 568 TPKMEKEGLEIMIGKKGIQGHNSCYLDSTLFCFLFATSSVLD-----TVLLRPK 618
Db 1923 GRP-IEDKQLEEDINKQ-----FDSVKLKEKPFASFKDYWEAYKAIGNIYELR 1970
QY 619 EKNDVEYSETQELLRTIENPLRIYGVYCATIMKLRKILEKVEAASGFTSEK---D 674
Db 1971 EVSD-----QERKMLKTARYTDPET--GKTVKEKIVGVNGLIFNNIQAAPAAQYVGRFN 2024
QY 675 PEE-----FLNLFHILRVEPLLKIRSAQKQVQDCYFYQIFMEKNEKVGVPITIOQLE 728
Db 2025 PEKNRVERIYENVYFLH---NPETNGRG-----FSKLPETIAVAFAHKLME 2066
QY 729 WSFINSNLKFAAPSCLIIQ--MPRFGKD 755
Db 2067 GAKIGNKTWIGLSNGLAGNIMEDYCKD 2095

RESULT 8
S38177
SSV7 protein homolog - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YKR098C
C:Species: Saccharomyces cerevisiae
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: S38177
R:Gaillon, L.; Dujon, B.
A:Submitted to the Protein Sequence Database, March 1994
A:Reference number: S38175
A:Accession: S38177
A:Molecule type: DNA
A:Residues: 1-717 <GAI>
A:Cross-references: UNIPROT:P36026; EMBL:Z28323; NID:G486592; PID:G486593; MIPS:YKR098C
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:UBP11
A:Cross-references: SGD:S0001806; MIPS:YKR098C
A:Map position: 11R

Query Match 2.6%; Score 130; DB 2; Length 717;
Best Local Similarity 23.9%; Pred. No. 0.85; Indels 126; Gaps 18;
Matches 84; Conservative 44; Mismatches 44

QY 392 SPPLQPPVNV--SLTTENRF---HSLPFLSLTKMPNTNGSIGHSPLSLAQSVMEELNTA 445
Db 138 SLPLRLPLVHMASLSLTHNKFDGSLHETPNELTKPTNDN-----SKEDIVRESNQI 187
QY 446 PVQESPLAMPNGNHSGLVGSIAEVKENPPFYGVIRWIGQPPGLNEVLAGELEDEBAG 505
Db 188 -----ASSNKLEAGS-----EV-----199
QY 506 CTDGTRGTRYFTC--ALKKALFVKLSKCRPSRPFASLQ-----PVSNQIERN 552
Db 200 -----AYTTSKEALSPPSYLKL--STGKDALFKTLSSPATAPPVHSLVSSQI--RDS 248
QY 553 SLAFGGYLSVVEENTPPKMEK---EGLEIMIGKK-----KGIQGHY 591

Db 249 SQDSSSLSKVEK-----PKEEKGKIEAESSAPKAYNLPVIEDSNLLSELITGLQNPC 304
QY 592 NSCYLSTDLTCLPAPFASVLDVLLRPKE--KNDVEYSETQELLRTIENPLRIY-----G 645
Db 305 NTCYINSIIICLLGFTTLFRDLFTLKRYRLFLNTNPKYKEVQ--LSRSIYVLFKKMGLNGR 363
QY 646 YVCATKIMKLRKILEKVEAASGFTSEKDEEPLNLL--PHHILRVEPLK 694
Db 364 AIIPNRLPKCK---KLRPDLNIPDQODTQOEPLIVLARIHEELSNNVVK 412

RESULT 9
T13827
kinesin-73 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T13827
R:Li, H.P.; Liu, Z.M.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 1086-1091, 1997
A:Title: Kinesin-73 in the nervous system of Drosophila embryos.
A:Reference number: Z17784; MUID:97188425; PMID:9037010
A:Accession: T13827
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1921 <LIH>
A:Cross-references: UNIPROT:O01349; EMBL:U81788; NID:G1906595; PID:G1906596; PIDN:AAB504
C:Genetics:
A:Cross-references: FlyBase:FBgn0019968

Query Match 2.5%; Score 125.5; DB 2; Length 1921;
Best Local Similarity 18.1%; Pred. No. 8.3;
Matches 120; Conservative 94; Mismatches 227; Indels 221; Gaps 30;

QY 27 ECSTVDKQTKLLKVPKSGIGQYIDRSVGHRSIPS---AKGKNQIGLKILEQPHAVLF 83
Db 1321 DCSASDGET-----YIEKYTRGVSAVESILTDLRLQNVAVKELETAHQPL 1367
QY 84 VDEDVVEINEKFTTELLAITNCEERFSLFKNRNRLSKGLQIDVGCVPKVLRSGE-----138
Db 1368 SMKRTSVV--FNFSQQLI-----NKLQIMRFDASWESLNVGRSESFADL 1411
QY 139 -----EKFP-----GVVFRGCPPLAERTVSGIFPGVELLEEGRCGGTGVYQ--181
Db 1412 NNSALGNKFTPAGHSPAGAGVIRSRHS-----FGKGSSDSFGKAF--GIASP 1459
QY 182 -----GKQLFOCDED--CGFVALDKLELIEDDTALESYAGPGDTMQVELPPELINSR 233
Db 1460 ATSKLJGMRMTLHEEPLGGHRSUDE---EPEDSYSDSEYAAEYQERQNKSWATRSR 1515
QY 234 VSLKGGETIESGTVIFCDVLPKGESLGYFVGVDMDNPIGNWDGRFDGVLCSFACVESTIL 293
Db 1516 --LTAKTMDS---FMDV--SSHNSQSYL-----SVTSSANAM 1547
QY 294 LHINDIPESVQERRPPKLAFMRSRGVGDKGSSSHN-----KPKATGSTD-----339
Db 1548 KHLTGLATLSMSSS-----TSSGYGSAVSCNNLSNEDIASMRSMIDETPDGFRVN 1599
QY 340 ---PGNRSEL---FYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSL---TELSTDF 389
Db 1600 SNSPPNQRARVNPFLKDKMPKAKIOEQPEQAKK---LQEAFTPLELESRENAQSDDE 1656
QY 390 RSSPPL-----OPPPVNSLTENRHFSLPFLTKM--PNTNGSIGHSPLSLAQ 437
Db 1657 CAQLPKNNNNVDLVNEKPLSGQMELEEPSQESQTEFATDQNGNRSDLSHSSD 1716
QY 438 VMEELNTAPVQESPL-----AMPNGNSHGLE-----464
Db 1717 LLE--GGIIVREELPAGKVVVRKKSNTQPPSPNGSNINNNNGTTQVPRINHRASVAKMEG 1774
QY 465 -----VGSIAEVKENP-----PF--YGVIRWIGQPPGLNEV 493
Db 1775 LAAYLDSSIMTSSTEVDEESKDELVLPEMLVWGESVLIRPVNTSGVIRFVGTTFBQPCA 1834

QY 494 LAGLEDEACACTGDTFRGTRYFTCAKALKFVKLKSRCRPSRPFASLOPVSNQIERCNS 553
Db 1835 WIGVEL-DTPPTGKNDGSKVGQYFOCKPKHGMFVRSRDKLMLDKRGKAMR-AYKAAEKSNS 1892
QY 554 LA 555
Db 1893 IS 1894

RESULT 10
F83853
S-methyltetrahydrofolate S-homocysteine methyltransferase meth [imported] - Bacillus hal
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: F83853
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F83853
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1146 <STO>
A:Cross-references: UNIPROT:Q9KCB1; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA053
A:Experimental source: strain C-125
C:Genetics:
C:Superfamily: cobalamin-dependent methionine synthase; cobalamin-binding homology

Query Match 2.5%; Score 124.5; DB 2; Length 1146;
Best Local Similarity 17.8%; Pred. No. 4.3;
Matches 153; Conservative 115; Mismatches 303; Indels 287; Gaps 40;
QY 52 DRSVGHSRIPSAKGGK--NOIGI-----KILEQPH-----AVLFVDEDVVEINEKF 95
Db 413 DEKVIEQALTYSGKAIINSINLEDGEERFEKVPLVHKYGAVVVGTIDEGNAT--- 469

QY 96 TELLAITNCERFSLFKNRRLSKG-----LQIDVGCVPVKQLRSGBEKFVGVRFRG 149
Db 470 AEKCLAV--AKRSYDLVKNYRIPSDIIFPLVFPVGTGDEQYIGSANETVEGIRRIKE 527

QY 150 PL---LAERTVGIFFGVLEBEGRGQFTDGVYQKQLFQCD-----DCGFVALDKLE-- 201
Db 528 ELPECLTILGVSNVSGLPVP---GREVLNAAV---LYHCTQAGLDYAVNTEKLERY 579

QY 202 -----LIEDDITAL---ESDYAGPGDTMQLPELPLEINRSVSLKGGTIES 244
Db 580 ASISDEKELSRKLLFETTDTELAEFTAFYRGKKAEEKVETSLTLEERLA-----NYIVE 635

QY 245 GTVIFCDVLPKESLGFGVVDMDNPIGNWGRFD---GVLCSFACVESTILLHNDIIP 301
Db 636 GS-----KDGL-----TEDLDKALAKYDDPLDIINGPLMN-GMDEVGRLENNELIV 681

QY 302 ESVTQERRPPKLAFMGRGVGDKSSSHNPKATGSTSDPGNRRSELYFTLNGSSVDSQPQ 361
Db 682 AEVLQSAEVM-----KASVAHLEPHMEKKADHCKGKI-ILATVKGD----- 722

QY 362 SSKSNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPPVNSLTENRPHSLPFLTKMPN 421
Db 723 -----VHDIGKGLVEI-----ILSNNGFRIVNIGIKVTSN 752

QY 422 -----TNGSIGHSPLSLSAQSVMEELNTAPVQESPPLAMPNGSHGLEVGSIA--- 469
Db 753 ELIEAVARENPDALGS--GLLVKSAQWVLTADQLKQQQISIP-----ILVGGALTR 804

QY 470 ---EVKENPPFYGVIRWIGOPPLNEVLALELEDECACTDGTFRGTRYFTCALKALF 526
Db 805 KFTNTKIAPYDGLVY-----AKOAMNGLAN----- 833

QY 527 VKLKSRCRPSR---FASLQPVSNQIERCNSLAFGG-----YLSEVVEENTPPKME 573
Db 834 ---KLMKPDREKLAVSLHEAKSQRTOQMGGGTAVAVKPTRSHVSTTVPVAVPPDVK 890

QY 574 KE-----GLEIMIGKKKIOGHYNSCYLDSTFLCLFAFSSVLDVTLRPKE 619
Db 891 PHILRHHSTAHLPEYINMQMLGRHLGLQCKVSRLLAEKDEKAL-ELKEKVDALLTRVKE 949
QY 620 K-----NDVEYVSET-----OELLRTEIVNP 640
Db 950 EQLMEAHGMVQFPFPAQSDGDDIVYDOTGTNETERFHPFQNKPEYVLCADFRLPVSSGE 1009
QY 641 LRIYGVYCATKIMKRLKILEKVEAASGFTSEKDPBEFLNLFHILRVPLKIRSAQ 700
Db 1010 MDVVGFLAVTAGKIGRELGEQAEADY-----LFSLHIQATALEMAEGFAE 1056
QY 701 KYQDCYFYQIFMEKNEKVGVP-----TIQOLLEWSFINSNLKFAEAPSCLIHQMPRGKD 755
Db 1057 RV-----HQLM---RDKWGFPDSADFTMEERFAAKYRGIRVSFG-YPAC-----PDLDDQ 1102

QY 756 FKLKKIFP-SLELNITD 772
Db 1103 AKLFKLLKPKGKIGIEUTE 1120

RESULT 11
AD0835
large repetitive protein [imported] - Salmonella enterica subsp. enterica serovar Typhi
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0835
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connerthon, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD0835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3624 <PAR>
A:Cross-references: GB:ALS13382; PIDN:CAD05867.1; PID:g16503842; GSPDB:GN00176
C:Genetics:
C:Gene: STY2875

Query Match 2.5%; Score 124; DB 2; Length 3624;
Best Local Similarity 21.9%; Pred. No. 29;
Matches 103; Conservative 61; Mismatches 167; Indels 140; Gaps 28;
QY 191 DCGFVALDKLELIEDDITALSDYAGPG---DTMQVELPPELNSRVSL-----KGGET 241
Db 1532 DSGLPVAPPAITAEDEDDVGSIQGNIAAGGATDDTMTPLRTGTTDGTSTVEVFDGDSAGPAT 1591

QY 242 IE-SGTVIFCDVLPKESLGFGVVDMDNPIGNWGRFD---GVLCSFACVESTILLHIN 297
Db 1592 VDASGNWIFFIATPLSESTHYFT-VQATN--ANGPGGLSAPVITVLSAPAQPVTSAT 1648

QY 298 DIIP-----ESVTOERRPPKLAFMGRGVGDKSSSH---NKPATGTSDD-PGNRRS 345
Db 1649 DDVPGMTGLDNGALINDSRP-----TLNGTGEAGATIRILDNGVEIGSATVDQSGNWR- 1702

QY 346 ELFYTLNGSSVDSQPSQSKSNWYIDEVAEDPAKSLTEISTDF---DRSSPPIQPPPVN 401
Db 1703 ---FTN-TELES-----NAHIFTAVATDPAGNSQGLSDGFTLNIDAQAPDV--PVIT 1749

QY 402 SLTTEPRFHSLLP-----SLTKMPTNGSIGHSPLSLSAQSVMEELNTAPVQESPPLAMP 456
Db 1750 SVDDNNNQPTVPVLPFGOSTDDROPILNGT-GEPCATITFDNGTPLGTAQGENGSWTFF 1808

QY 457 -PGN---SHGLEV-----GSLAEVKE-----NPPFYGV I-RWIGOPPLNEVLAG 496
Db 1809 VPRNLSEGSNLTVSATDPAGNTSAVAPWTIVVDITFPAIPVLTSDVDDQDQGITGLVS 1868

QY 497 LELEDEC-----AGCT-----DGTF-----RGTRYFTC 519
Db 497 LELEDEC-----AGCT-----DGTF-----RGTRYFTC 519

Db 1869 GOLTNDATPLNGRGEAGATINVYLDGNPASIGTTTNSDGTWSFTPTQPLANGSHFTT- 1927

QY 520 ALKKALFVKLSCRPSRPFASIQPVSNQIQRNCNSLAFGGYLSVVBENTPP 570

Db 1928 -----LSATDPAGN-----SSAVSSGFVLTII--DTTPP 1953

RESULT 12

S63244

BN11 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein N0646; protein YNL271c

C:Species: Saccharomyces cerevisiae

C:Date: 27-Apr-1996 #sequence revision 03-May-1996 #text_change 09-Jul-2004

C:Accession: S63244; S63245; S48523; S60909; S65111

R:Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.

A:Submitted to the Protein Sequence Database, April 1996

A:Reference number: S63235

A:Accession: S63245

A:Molecule type: DNA

A:Residues: 1-1553 <SEN>

A:Cross-references: UNIPROT:P41832; EMBL:Z71547; MIPS:YNL271c

A:Experimental source: strain S288C

R:Messeguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.

A:Submitted to the Protein Sequence Database, April 1996

A:Reference number: S63245

A:Accession: S63245

A:Molecule type: DNA

A:Residues: 987-1953 <MES>

A:Cross-references: EMBL:Z71547; MIPS:YNL271c

A:Experimental source: strain S288C

R:Fares, H.F.; Pringle, J.R.

A:Submitted to the EMBL Data Library, April 1994

A:Description: Synthetic Lethals of CDC12.

A:Reference number: S48524

A:Accession: S48523

A:Molecule type: DNA

A:Residues: 1-937, 'A', 939-1429, 'C', 1431-1953 <FAR>

A:Cross-references: EMBL:L31766; NID:G472524; PIDN:AAA34455.1; PID:G472525

R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.

A:Submitted to the EMBL Data Library, October 1995

A:Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV from Sa

A:Reference number: S60909

A:Accession: S60909

A:Molecule type: DNA

A:Residues: 1-1553 <SE2>

A:Cross-references: EMBL:X92494; NID:G1045236; PIDN:CAA63225.1; PID:G1045237

R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.

Yeast 12, 505-514, 1996

A:Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa

A:Reference number: S65111; MUID:96310631; PMID:8740425

A:Accession: S65111

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1553 <SEW>

A:Cross-references: EMBL:X92494; NID:G1045236; PIDN:CAA63225.1; PID:G1045237

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: SGD:BN11; SHE5; SYL39

A:Cross-references: SGD:S0005215; MIPS:YNL271c

A:Map position: 14L

Query Match

Best Local Similarity 19.3%; Pred. No. 14;

Matches 181; Conservative 100; Mismatches 300; Indels 357; Gaps 46;

QY 57 HGRIPSAKGNQIGLKLEOPHAVLFVDEDDVVEINE--KFTLELLAIT-----NCEER 108

Db 902 HIKPKVKQKQKQK-DRKPLVKPQ-----EADVKNLDRALTEIQMESNDISKFNVEER 955

QY 109 FS-LFKNRNRLSKGLQIDQVCPVKVQKVRGSRGEEKPPGVPFRGPELLAERTVSGIFFGVYELL 167

Db 956 VNELFNEKSLA-----LKRLEKLETKYKGF-----GIDFNVDIEI 990

QY 168 EEGRGQGTDDVYQKQLFOCEDDCGFVALD-----KL 200

Db 991 MDSPKNTGD-----VETESDANYASLDPTYQKKLDEINRITDOLLDIQTQTEHEI 1042

QY 201 ELTEDDDTALESYAG--PGDTMQVELPPLLEINRSVS-LKGGETIESGTVIFCDVLPQKE 257

Db 1043 QVEDGESDLSSSSDDESEIYQDASPTQELRSEHSELSSG-----SGPGSFLDALSQKY 1098

QY 258 SLGYFVGVDV---DNPIGNWGRFPGVLCFACVESTILLHINDIIPESVTOERRPPKLA 314

Db 1099 GTGQNVTAASAAFGNNNGS-----GIGPLHKSVEKTFM---NRLRSTVS---APYLE 1146

QY 315 FMGRGVG-----DKGSSSHNKPKATGSTSDPGNRRSELFTYLANG----- 353

Db 1147 ELTKQNVKPEYQNEDEGLDKKSLPENSTASAAAFDKAKMRQ--HVENKQGRVVNH 1205

QY 354 ---SSVDSQPQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPV-----NS 402

Db 1206 EEDKTADFSAVSKLNT-----DGAEDLSQSSVLSQPPPPPPPPVPAKLFGES 1257

QY 403 LTTENRPHSLPFLTKMPTNTSGHSPLSLSAQSVMEELNTPVQESPPPLAMPPGNSHG 462

Db 1258 LEKEKKSSED--DTVKQETT-----GDSP-----APPPPPPPPPPPMALFG 1296

QY 463 LEVGSALAEVKNPFYGVIRWIGOPPCGLNEVLAG----- 496

Db 1297 KPKG-----ETPP-----PPPLPSVLSSTDGVIPPPAPPMPMPASQIKSAVTSP 1340

QY 497 -----LEDEBCAGCTDGTFRGT-----RYFT 518

Db 1341 LPQSPSLFEKYPRPHKKLKQHWKCLDCTNSIWTGKAEFADDLYEKGVLADLEKAPA 1400

QY 519 CALKXALFVKLSCRPSRPFASIQPVSNQ-----IERCNSLAFGGYLSVSVW-ENTP 569

Db 1401 AREIKSLASKRKEDLQKITFLS-RDISQQGGINLHMYSSLSVADLVKILNCNDRDRLQTP 1459

QY 570 PKME-----KEGLEIMTKKGIQGHYNSCYLDSTLFCLEAFSSVDTVLLRPKEK--NDV 623

Db 1460 SVVEFLSKSIIIEVSV---NLARNYAPYSTD-----WEGVRNLEDAKPEKDPNDL 1507

QY 624 E-----YYSTQELLRTIEINPLRIYGVCAATKIMKILKIEKVEA---- 664

Db 1508 QRADQIYQLMVNLESYWGSRMRAL-TVVTYSEREYNELLA---KLKVKDKAVSALOES 1562

QY 665 -----ASG-----FTSEKDPPEEFLNIL----- 682

Db 1563 DNRNVNVLAVGNFNDTSKQAGFKSLTQLRLTFIKDTNSMTFLNYVEKIVRLNYP 1622

QY 683 -FHILL-RVEPLLKI--RSAGQKQVDCYFYQIFMEKNEKVGVTPTIQQLLEWSFINSLKP 738

Db 1623 SFNDFLSELEPVDVVKVSIQVLVNDV-----KQFSQSVNVNRSVEIGNLSDSKF 1674

QY 739 AEAPSLIIQMPRFGDKFLFKIFPSLELNITDLED 776

Db 1675 HPLDKVLIKTLP-----VLPEARKK-GDLED 1700

RESULT 13

T02345

hypothetical protein KIAA0324 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004

C:Accession: T02345

R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.; re, J.; White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.

submitted to the EMBL Data Library, March 1998

A:Description: Sequencing of human chromosome 16p13.3.

A:Reference number: Z14664

A:Accession: T02345

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1791 <RIC>

A:Cross-references: UNIPROT:O60382; EMBL:AC004493; NID:g2996648; PIDN:AA08453.1; PID:g2

308 SDICTSLPQNSQIQOMHPQNIQS--DSNSNSFNLAUVGVKSESSPOGWPSKQENTLMS 366
 421 NTNGSIGH-----SPLSLSAQSYMEELNAPVQESPPLAMPNGNSHGLE- 464
 367 NAISSGKHQIEDPRQITGWDEAQPNNLTGSGVIGQHTSTISE-----SHNLQN 416
 465 -VGS---LAEVKENPPFYGVIRWI-----GPPGLNEVLAGLEHLEDECA- 504
 417 SIGTTCRYGNVSHDPFKQKQORWLLFLRHARSKCPGP-----GRCDQNCVTVQKLWSH 470
 505 --GCTDGTGRGTRYFYCALKKALKFVKLKSCRDPDSFASLQPVSNQIETRCN----- 552
 471 MDNCAPOQ---CLYPRCRHTKALIGHYKNCK-DPRCPVCVPVTVYQQOANRALARKNE 526
 553 SLAFGGYLSEWVE-----ENTPPKMEKEGLEIMTKKKGIQGHY 591
 527 SSAVGSVNRSVVNSDLSANAGAVSGTPRCADTLNQLPSLKRLLKVE-----QSFQP-- 578
 592 NSCYLDSTUFLPAFSSVLDTVLLRPEKNDVE-YYSETQELRLTEIVNPLRIYGVYCAT 655
 579 ---VWPKTESCKSSIVSTTEADLSQDAERKDRHPLKSETME-VKVEIPD----- 623
 651 KIMKLKILEKVEAASGFTSEBKDPFEFLNLFPHHILRVEPLKIRSAQKVQDCYFYQI 710
 624 -----NSVQAGFIKETSEP-----FENVPKPKVPSEPGKHGSGDSPKQENI 667
 711 FMEK-----NEKVGVTPTIOQLLEWSFINSNLKFAEAPSCLIIQMPRFGKDFKLFKTFPS 765
 668 KMKKEGWPKEGPGCPKBELVSPELTS-----KSRKPK 702
 766 LE-LNITDLEDTPRQCR-----ICGGLAMVECREC-----Y 796
 703 IKGVSLTELF--TPEQVREHIRGLRWVGQSKAKAEKNQAWMSMSNSCOLCAVEKLTF 760
 797 DDPDISAGKIKQCKCTCNTQVHLHPKRLNHYNPVSLPKDLPDWDWRH-CCIPCONMELF 855
 761 EPPPI-----YCTPCGARI-----KRNAMYTVG-----GGETHYFCIPCYNSSRG 802
 856 AVLCLIE-TSHYVAFVKYKDD---SAWLFFDSM-----ADRDGGQGNFIP 897
 803 DTILAEGTSPKAKLEKKKNDIEIESWVQCDKQAWHQICAFNGRRNDGGQAEYTCP 862
 RESULT 15
 T10955
 early nodulin binding protein 1 - spring vetch
 C;Species: *Vicia sativa* (spring vetch, tare)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C;Accession: T10955
 R;Christiansen, A.; Hansen, A.C.; Vijn, I.; Pallisgard, N.; Larsen, K.; Yang, W.
 submitted to the EMBL Data Library, December 1995
 A;Description: A novel type of DNA binding protein interacts with a conserved s
 A;Reference number: Z17228
 A;Accession: T10955
 A;Status: preliminary;
 A;Molecule type: mRNA
 A;Residues: 1-1641<CHR>
 A;Cross-references: UNIPROT:Q41700; EMBL:X95995; NID:gl360633; PID:e225826
 C;Genetics:
 A;Note: ENBP1
 C;Keywords: DNA binding
 Query Match 2.4%; Score 121; DB 2; Length 1641;
 Best Local Similarity 19.4%; Pred. No. 13;
 Matches 210; Conservative 109; Mismatches 311; Indels 452; Gaps 57
 Qy 107 ERFSLFKNNRLSKGLQIDVGCVPKVQLRSGE-----EKFPG-VVFRFGPLLAERTVSGI 160
 Db 100 ESFNLMQE---GEGQQVEV-----FEGSGNLKGLFDGVEFLGGFVEDRNVVGL 146
 Qy 161 -----FFGVELLEBGRGGQFTDGVYQKQLFCQDEDCGFVALDKLELIEDDDTALES 212
 Db 147 GOPSSVGVFG-----NAGGVAGVGKEDHGSRVDGVCQ-----NDSLAPH 187

Search completed: April 18, 2005, 15:25:28
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 15:22:19 ; Search time 355 Seconds
(without alignments)
3122.360 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCWQSPMTSLYK 949

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US097A_COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US105_COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US106_COMB.pcp.*
- 33: /cgn2_6/ptodata/1/paa/US107_COMB.pcp.*
- 34: /cgn2_6/ptodata/1/paa/US108_COMB.pcp.*
- 35: /cgn2_6/ptodata/1/paa/US109_COMB.pcp.*
- 36: /cgn2_6/ptodata/1/paa/US110_COMB.pcp.*
- 37: /cgn2_6/ptodata/1/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT 1
US-09-671-687A-3
; Sequence 3, Appli
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: CANTARELLA, Giuseppe
; TITLE OF INVENTION: INHIBITOR OF NF-KB ACTIVATOR
; FILE REFERENCE: WALLACH=25
; CURRENT APPLICATION NUMBER: US/09/671,687A
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/646,403
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: IL 126024
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: IL 134604
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1

ALIGNMENTS

Result No.	Score	Query Match	Length	DB	ID	Description
1	5034	100.0	949	20	US-09-671-687A-3	Sequence 3, Appli
2	4983	99.0	953	1	PCT-US02-27777-137	Sequence 137, App
3	4983	99.0	953	1	PCT-US02-27777-137	Sequence 137, App
4	4983	99.0	953	27	US-10-170-205E-37570	Sequence 37570, A
5	4983	99.0	960	1	PCT-US04-07268-250	Sequence 250, App
6	4983	99.0	960	33	US-10-788-792-250	Sequence 250, App
7	4971.5	98.8	956	1	PCT-US02-14570-4	Sequence 4, Appli
8	4971.5	98.8	956	23	US-09-851-673-4	Sequence 4, Appli
9	4971.5	98.8	956	33	US-10-755-889-490	Sequence 490, App
10	4971.5	98.8	956	37	US-60-440-068-490	Sequence 490, App
11	4971.5	98.8	956	37	US-60-469-757-490	Sequence 490, App
12	3876.5	77.0	739	1	PCT-US01-01239-1743	Sequence 1743, Ap
13	3876.5	77.0	739	22	US-09-764-902-1743	Sequence 1743, Ap
14	3862	76.7	731	22	US-09-786-797B-9	Sequence 9, Appli
15	3862	76.7	731	35	US-10-921-707-9	Sequence 9, Appli
16	3862	76.7	731	37	US-60-131-321-7	Sequence 7, Appli
17	3855	76.6	731	20	US-09-629-469A-18843	Sequence 18843, A
18	3855	76.6	731	35	US-10-917-503-18843	Sequence 18843, A
19	3636	72.2	685	18	US-09-488-725A-2399	Sequence 2399, Ap
20	3636	72.2	685	28	US-10-258-898A-2399	Sequence 2399, Ap
21	3636	72.2	685	28	US-10-286-897-2399	Sequence 2399, Ap
22	3563	70.8	698	18	US-09-488-725A-5971	Sequence 5971, Ap
23	3563	70.8	698	28	US-10-258-898A-5971	Sequence 5971, Ap
24	3563	70.8	698	28	US-10-286-897-5971	Sequence 5971, Ap
25	2755	54.7	558	1	PCT-US01-01239-1207	Sequence 1207, Ap
26	2755	54.7	558	22	US-09-764-902-1207	Sequence 1207, Ap
27	2455	48.8	512	1	PCT-US01-14827-12400	Sequence 12400, A
28	2343	46.5	476	20	US-09-629-469A-18587	Sequence 18587, A
29	2343	46.5	476	35	US-10-917-503-18587	Sequence 18587, A
30	1751	34.8	394	1	PCT-US01-14827-12399	Sequence 12399, A
31	1275	25.3	261	1	PCT-US01-16450-1609	Sequence 1609, Ap
32	1275	25.3	261	1	PCT-US01-16450A-1609	Sequence 1609, Ap
33	1275	25.3	261	28	US-10-264-237-1609	Sequence 1610, Ap
34	1173	23.3	238	1	PCT-US01-16450A-1610	Sequence 1610, Ap
35	1173	23.3	238	28	US-10-264-237-1610	Sequence 1610, Ap
36	1173	23.3	238	28	US-09-614-150-11799	Sequence 11799, A
37	856	17.0	517	20	US-09-614-150A-11799	Sequence 11799, A
38	856	17.0	517	37	US-60-191-637-11831	Sequence 11831, A
39	856	17.0	517	37	US-60-191-681-9275	Sequence 9275, Ap
40	786.5	15.6	167	1	PCT-US01-14827-15309	Sequence 15309, A
41	775	15.4	188	22	US-09-758-472-8763	Sequence 8763, Ap
42	775	15.4	188	28	US-10-235-926-8763	Sequence 8763, Ap
43	742.5	14.7	583	37	US-60-167-217-11813	Sequence 11813, A
44	742.5	14.7	583	37	US-60-173-464-9570	Sequence 9570, Ap
45	742.5	14.7	583	37	US-60-173-464-9570	Sequence 9570, Ap

; SEQ ID NO 3
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-687A-3

Query Match 100.0%; Score 5034; DB 20; Length 949;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 949; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSGLWSQEKVTSFYWEERIFVLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHSRI 60
DB 1 MSSGLWSQEKVTSFYWEERIFVLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHSRI 60
QY 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLLAITNCERFSLFKNRNLK 120
DB 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLLAITNCERFSLFKNRNLK 120
QY 121 GLQIDVGCPCVKQLRSGEKFPVGRFGRPLLAERTVSGIFFGVLELLEGRGGQFTDGVY 180
DB 121 GLQIDVGCPCVKQLRSGEKFPVGRFGRPLLAERTVSGIFFGVLELLEGRGGQFTDGVY 180
QY 181 QGKQLFOCEDDCGFVALDKLELIEDDDTALESYAGPGDTMQVELPPLINRSVSLKGE 240
DB 181 QGKQLFOCEDDCGFVALDKLELIEDDDTALESYAGPGDTMQVELPPLINRSVSLKGE 240
QY 241 TTIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHNDII 300
DB 241 TTIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHNDII 300
QY 301 PESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNRRSELFTYLANGSSVDSQP 360
DB 301 PESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNRRSELFTYLANGSSVDSQP 360
QY 361 QSKSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNNRHSLSLPSLTMP 420
DB 361 QSKSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNNRHSLSLPSLTMP 420
QY 421 NTNGSIHSPSLSAOSVMEELNTAPVQSPPLAMPNGSHGLEVSLAEVKNPFPYGV 480
DB 421 NTNGSIHSPSLSAOSVMEELNTAPVQSPPLAMPNGSHGLEVSLAEVKNPFPYGV 480
QY 481 IRWIGOPPGNEVLAGELEDECACTDGTFRGTRYFTCALKKALFKVLSKCRPDSRFAS 540
DB 481 IRWIGOPPGNEVLAGELEDECACTDGTFRGTRYFTCALKKALFKVLSKCRPDSRFAS 540
QY 541 LQVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600
DB 541 LQVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600
QY 601 FCLFAFSSVLDVLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVYCATKIMKLKILE 660
DB 601 FCLFAFSSVLDVLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVYCATKIMKLKILE 660
QY 661 KVEAASGFTSEEDKDPREFNLPHILRVBPLLKIRSAGOKVQDCYFYQIFMEKNEKVG 720
DB 661 KVEAASGFTSEEDKDPREFNLPHILRVBPLLKIRSAGOKVQDCYFYQIFMEKNEKVG 720
QY 721 PTIQQLLEWFSFNSNLKFAEAPSLIIQMPRFGDKPKLKKIPPSLELNTDLEETPRQ 780
DB 721 PTIQQLLEWFSFNSNLKFAEAPSLIIQMPRFGDKPKLKKIPPSLELNTDLEETPRQ 780
QY 781 CRICGGLAMEYECRECYDDPDISAGKIQPKCTNTQVHLHPKRLNHNKYNVSLPKOLDW 840
DB 781 CRICGGLAMEYECRECYDDPDISAGKIQPKCTNTQVHLHPKRLNHNKYNVSLPKOLDW 840
QY 841 DWRHGICPCQNMELFAVLICETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 900
DB 841 DWRHGICPCQNMELFAVLICETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 900
QY 901 PCPEVGEYLKWSLEDLHLSRSRIQGCARLLCDAYNCMTQSPMSLYK 949
DB 901 PCPEVGEYLKWSLEDLHLSRSRIQGCARLLCDAYNCMTQSPMSLYK 949

RESULT 2

PCT-US02-27777-137
; Sequence 137, Application PC/TUS0227777
; GENERAL INFORMATION:

; APPLICANT: diaDexus, Inc.

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; APPLICANT: Salceda, Susana

; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P

; FILE REFERENCE: DEX-0346

; CURRENT APPLICATION NUMBER: PCT/US02/27777

; CURRENT FILING DATE: 2002-08-29

; PRIOR APPLICATION NUMBER: US 60/316,306

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 170

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 137

; LENGTH: 953

; TYPE: PRT

; ORGANISM: Homo sapien

PCT-US02-27777-137

Query Match 99.0%; Score 4983; DB 1; Length 953;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSFYWEERIFVLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHSRI 60
DB 1 MSSGLWSQEKVTSFYWEERIFVLLLOECSTVDKOTKLLKVPKSGIGQYIQDRSVGHSRI 60
QY 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLLAITNCERFSLFKNRNLK 119
DB 61 PSAKGKKNQIGLKILQPHAVLFVDEDDVVEINEKFTTELLLAITNCERFSLFKNRNLK 120
QY 120 KGLQIDVGCPCVKQLRSGEKFPVGRFGRPLLAERTVSGIFFGVLELLEGRGGQFTDGV 179
DB 120 KGLQIDVGCPCVKQLRSGEKFPVGRFGRPLLAERTVSGIFFGVLELLEGRGGQFTDGV 180
QY 180 YQKQQLFOCEDDCGFVALDKLELIEDDDTALESYAGPGDTMQVELPPLINRSVSLK 238
DB 180 YQKQQLFOCEDDCGFVALDKLELIEDDDTALESYAGPGDTMQVELPPLINRSVSLK 240
QY 239 GETIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHIN 297
DB 241 GETIESGTVIFCDVLPKESLGYPVGVDMNDPIGNWDRFGDVLCSFACVESTILLHIN 300
QY 298 DIIPESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNR-RSELFYTLNGSSV 356
DB 301 DIIPESVTQERRPPKLAFMRSRGVGDGSSSHNKPATGSTSDPGNRSELFTYLANGSSV 360
QY 357 DSQPSQSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNNRHSLSLPSL 416
DB 361 DSQPSQSKNTWTIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSLTNNRHSLSLPSL 420
QY 417 TKMPNTNGSIHSPSLSAOSVMEELNTAPVQSPPLAMPNGSHGLEVSLAEVKNPFPY 476
DB 421 TRMPNTNGSIHSPSLSAOSVMEELNTAPVQSPPLAMPNGSHGLEVSLAEVKNPFPY 480
QY 477 FYGVIRWIGOPPGNEVLAGELEDECACTDGTFRGTRYFTCALKKALFKVLSKCRPDS 536
DB 481 FYGVIRWIGOPPGNEVLAGELEDECACTDGTFRGTRYFTCALKKALFKVLSKCRPDS 540
QY 537 RFASLQPVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
DB 541 RFASLQPVSNQIERCNSLAFGGVYLSEVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 600
QY 597 DSTLFCFLFAFSSVLDVLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVYCATKIMKL 656
DB 601 DSTLFCFLFAFSSVLDVLLRPKEKNDVEYSETQELLRTTEIVNPLRIYGVYCATKIMKL 660
QY 657 KILEKVEAASGFTSEEDKDPREFNLPHILRVBPLLKIRSAGOKVQDCYFYQIFMEKNE 716

Db 661 KILEKVEAASGFTSEBKDPSEFNILPHHLLRVEPLKIRSAQKQVQDCYFYQIFMEKNE 720
QY 717 KVGVPPIQQLLEWSFINSNLKFAEAPSCIIQMPRFCKDFKPKKIPPSLELNITDLED 776
Db 721 KVGVPPIQQLLEWSFINSNLKFAEAPSCIIQMPRFCKDFKPKKIPPSLELNITDLED 780
QY 777 TPRQCRICGGLAMYECCYDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVSPLPKD 836
Db 781 TPRQCRICGGLAMYECCYDDPDISAGKIKQFCKTQNTQVHLHPKRLNHNKYNPVSPLPKD 840
QY 837 LPDWDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKQDSDAWLFFDSMADRDGQNGFNI 896
Db 841 LPDWDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKQDSDAWLFFDSMADRDGQNGFNI 900
QY 897 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 949
Db 901 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSPMTSLYK 953

RESULT 3

PCT-US02-27777A-137

; Sequence 137, Application PC/TUS0227777A

; GENERAL INFORMATION:

; APPLICANT: diadexus, Inc.

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; APPLICANT: Salceda, Susana

; TITLE OF INVENTION: Compositions and Methods related to Breast Specific Genes and P

; FILE REFERENCE: DEX-0346

; CURRENT APPLICATION NUMBER: PCT/US02/27777A

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: US 60/316,307

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 170

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 137

; LENGTH: 953

; TYPE: PRT

; ORGANISM: Homo sapien

PCT-US02-27777A-137

Query Match 99.0%; Score 4983; DB 1; Length 953;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSVTDKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
Db 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSVTDKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
QY 61 PSAKGNKQIGLKILEQPHAVLFDVE-DVVEINEKETELLAITNCEERFSLPKRNRLS 119
Db 61 PSAKGNKQIGLKILEQPHAVLFDVEKDVVEINEKETELLAITNCEERFSLPKRNRLS 120
QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 179
Db 121 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 180
QY 180 YQKQLFQDCEDCG-FVALDKLELIEDDDTALESYAGPDTMQVLPPLPINSRVSLKG 238
Db 181 YQKQLFQDCEDCGFVALDKLELIEDDDTALESYAGPDTMQVLPPLPINSRVSLKV 240
QY 239 GETIESGTIVFCQVLPKESLSGYFVGVDMDNPIGNWDGRPDGV-LCSFACVESTILLHN 297
Db 241 GETIESGTIVFCQVLPKESLSGYFVGVDMDNPIGNWDGRPDGVQVLCFACVESTILLHN 300
QY 298 DIIPESVTOERRPPKLAFWRSRGVGDGKSSHNKPKATGSTDSPGNR-RSELFTYLANGSSV 356
Db 301 DIIPESVTOERRPPKLAFWRSRGVGDGKSSHNKPKATGSTDSPGNRNRSELFTYLANGSSV 360
QY 357 DSQPSKSKNTWYIDVAVDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRPHSLPFSL 416
Db 361 DSQPSKSKNTWYIDVAVDPAKSLTEISTDFDRSSPPLQPPVNSLTTENRPHSLPFSL 420

} late

RESULT 4

US-10-170-205E-37570

; Sequence 37570, Application US/10170205E

; GENERAL INFORMATION:

; APPLICANT: ADAMS, Mark

; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN

; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF

; FILE REFERENCE: CL001381

; CURRENT APPLICATION NUMBER: US/10/170, 205E

; CURRENT FILING DATE: 2002-06-13

; NUMBER OF SEQ ID NOS: 40312

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 37570

; LENGTH: 953

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-170-205E-37570

Query Match 99.0%; Score 4983; DB 27; Length 953;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSVTDKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
Db 1 MSSGLWSQEKVTSPLYWEERIFYLLLOECSVTDKOTKLLKVPKGSIGQYIQDRSVGHSRI 60
QY 61 PSAKGNKQIGLKILEQPHAVLFDVE-DVVEINEKETELLAITNCEERFSLPKRNRLS 119
Db 61 PSAKGNKQIGLKILEQPHAVLFDVEKDVVEINEKETELLAITNCEERFSLPKRNRLS 120
QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 179
Db 121 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 180
QY 180 YQKQLFQDCEDCG-FVALDKLELIEDDDTALESYAGPDTMQVLPPLPINSRVSLKG 238

} late

Db 181 YQKQLFQDCEGCVFVLDKLEIEDDDTALSDYAGPDTMQVELPLEINSRSLKV 240
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDGRFDGV-LCSFACVESTILLHIN 297
Db 241 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDGRFDGV-LCSFACVESTILLHIN 300
Qy 298 DIIPESVTOERRPPKLAFAFMSRGVGDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 356
Db 301 DIIPESVTOERRPPKLAFAFMSRGVGDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 360
Qy 357 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFRSPPLPPVNSLTITTENRPHSLPFSL 416
Db 361 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFRSPPLPPVNSLTITTENRPHSLPFSL 420
Qy 417 TKMPNTNGSIGHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEVKNPP 476
Db 421 TKMPNTNGSIGHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEVKNPP 480
Qy 477 FYGVIRWIGQPPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKKALFVKLKS CRPDS 536
Db 481 FYGVIRWIGQPPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKKALFVKLKS CRPDS 540
Qy 537 RFASLOPVSNQIBRCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
Db 541 RFASLOPVSNQIBRCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 600
Qy 597 DSTLFCFLAFSSVLDVTLRPRKKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 656
Db 601 DSTLFCFLAFSSVLDVTLRPRKKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 660
Qy 657 KILEKVEAASGFTSEBKDPEEFNLILFHILRLRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
Db 661 KILEKVEAASGFTSEBKDPEEFNLILFHILRLRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 720
Qy 717 KVGVPITQOLLEWSFINSNLKFAEAPSLIIQMPRFKDFKFKIIPPSLELNITDLED 776
Db 721 KVGVPITQOLLEWSFINSNLKFAEAPSLIIQMPRFKDFKFKIIPPSLELNITDLED 780
Qy 777 TPRQCRICGLAMVYECRECYDDPDISAGKIQKCTCNTQVHLHPKRLNHNKYNPVS LPKD 836
Db 781 TPRQCRICGLAMVYECRECYDDPDISAGKIQKCTCNTQVHLHPKRLNHNKYNPVS LPKD 840
Qy 837 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLFDDSMADRDGGQGNFI 896
Db 841 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLFDDSMADRDGGQGNFI 900
Qy 897 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 949
Db 901 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 953

RESULT 5
PCT-US04-07268-250
; Sequence 250, Application PC/TUS0407268
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: PCT/US04/07268 } late
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 250
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-07268-250

Query Match 99.0%; Score 4983; DB 1; Length 960;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;
Qy 1 MSSGLWSQKVTSPYWEERIFYLLLOBCSVTDKQTKLLKVPKSGIQYIQDRSVGHSRI 60
Db 8 MSSGLWSQKVTSPYWEERIFYLLLOBCSVTDKQTKLLKVPKSGIQYIQDRSVGHSRI 67
Qy 61 PSAKGKKNQIGLKILEQPHAVLFVDE-DVVEINEKEFTTELLAINTNCEERFSLFKNRRLS 119
Db 68 PSAKGKKNQIGLKILEQPHAVLFVDEKDVVEINEKEFTTELLAINTNCEERFSLFKNRRLS 127
Qy 120 KGLQIDVGCPCVKQLRSBEEKFPGVVRFRGPLLAERTVSGIFFGVELLEBGRGGFTDGV 179
Db 128 KGLQIDVGCPCVKQLRSBEEKFPGVVRFRGPLLAERTVSGIFFGVELLEBGRGGFTDGV 187
Qy 180 YQKQLFQDCEGCVFVLDKLEIEDDDTALSDYAGPDTMQVELPLEINSRSLKV 238
Db 188 YQKQLFQDCEGCVFVLDKLEIEDDDTALSDYAGPDTMQVELPLEINSRSLKV 247
Qy 239 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDGRFDGV-LCSFACVESTILLHIN 297
Db 248 GETIESGTIVFCVLPKESLGYFVGVDMNDPIGNWDGRFDGV-LCSFACVESTILLHIN 307
Qy 298 DIIPESVTOERRPPKLAFAFMSRGVGDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 356
Db 308 DIIPESVTOERRPPKLAFAFMSRGVGDGKSSSHNPKATGSTDGPNR-RSELYFTLNGSSV 367
Qy 357 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFRSPPLPPVNSLTITTENRPHSLPFSL 416
Db 368 DSQPSKSNWTYIDEVAEDPAKSLTEISTDFRSPPLPPVNSLTITTENRPHSLPFSL 427
Qy 417 TKMPNTNGSIGHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEVKNPP 476
Db 428 TKMPNTNGSIGHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSLAEVKNPP 487
Qy 477 FYGVIRWIGQPPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKKALFVKLKS CRPDS 536
Db 488 FYGVIRWIGQPPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKKALFVKLKS CRPDS 547
Qy 537 RFASLOPVSNQIBRCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 596
Db 548 RFASLOPVSNQIBRCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYL 607
Qy 597 DSTLFCFLAFSSVLDVTLRPRKKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 656
Db 608 DSTLFCFLAFSSVLDVTLRPRKKNDEYVYSETQELLRTIENVPLRIYGVVCATKIMKLR 667
Qy 657 KILEKVEAASGFTSEBKDPEEFNLILFHILRLRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
Db 668 KILEKVEAASGFTSEBKDPEEFNLILFHILRLRVEPLLKIRSAQKQVQDCYFYQIFMEKNE 727
Qy 717 KVGVPITQOLLEWSFINSNLKFAEAPSLIIQMPRFKDFKFKIIPPSLELNITDLED 776
Db 728 KVGVPITQOLLEWSFINSNLKFAEAPSLIIQMPRFKDFKFKIIPPSLELNITDLED 787
Qy 777 TPRQCRICGLAMVYECRECYDDPDISAGKIQKCTCNTQVHLHPKRLNHNKYNPVS LPKD 836
Db 788 TPRQCRICGLAMVYECRECYDDPDISAGKIQKCTCNTQVHLHPKRLNHNKYNPVS LPKD 847
Qy 837 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLFDDSMADRDGGQGNFI 896
Db 848 LPDMDWRHGCIPCONNELFAVLCTIETSHYVAFVKYKDDSAWLFDDSMADRDGGQGNFI 907
Qy 897 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 949
Db 908 PQVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK 960

RESULT 6
US-10-788-792-250
; Sequence 250, Application US/10788792
; GENERAL INFORMATION:

; APPLICANT: Bayer Pharmaceuticals Corporation
 ; APPLICANT: Eveleigh, Deepa
 ; APPLICANT: Bigwood, Douglas
 ; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
 ; FILE REFERENCE: 5152
 ; CURRENT APPLICATION NUMBER: US/10/788,792
 ; PRIOR FILING DATE: 2004-02-27
 ; PRIOR APPLICATION NUMBER: US 60/450,655
 ; PRIOR FILING DATE: 2003-02-28
 ; NUMBER OF SEQ ID NOS: 254
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 250
 ; LENGTH: 960
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-788-792-250

Query Match 99.0%; Score 4983; DB 33; Length 960;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 948; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 1 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 60
 DB 8 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 67
 QY 61 PSAKKNQIGLKILEQPHAVLFVDE-DVVEINEKFTTELLAITNCEERFSLFKNRNLS 119
 DB 68 PSAKKNQIGLKILEQPHAVLFVDEKDVVEINEKFTTELLAITNCEERFSLFKNRNLS 127
 QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 179
 DB 128 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 187
 QY 180 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPLEINSRVSXKG 238
 DB 188 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPLEINSRVSXKG 247
 QY 239 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHN 297
 DB 248 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHN 307
 QY 298 DIIPESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNR-RSELPFTYTLNGSV 356
 DB 308 DIIPESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNR-RSELPFTYTLNGSV 367
 QY 357 DSQPSKSNVTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSLPFS 416
 DB 368 DSQPSKSNVTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSLPFS 427
 QY 417 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMP-PGNSHGLEVGSIAEVKENPP 476
 DB 428 TKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMP-PGNSHGLEVGSIAEVKENPP 487
 QY 477 FYGVIRWIOQPPGLNEVLVAGLEDEBCAGCTDGTFRGTRYFTCALKKALFVKLKSCRPS 536
 DB 488 FYGVIRWIOQPPGLNEVLVAGLEDEBCAGCTDGTFRGTRYFTCALKKALFVKLKSCRPS 547
 QY 537 REASLPQVSNQIERNCSLAFGGVLSWEENTPPKMEKEGLEIMICKKGIQGHYNSCYL 596
 DB 548 REASLPQVSNQIERNCSLAFGGVLSWEENTPPKMEKEGLEIMICKKGIQGHYNSCYL 607
 QY 597 DSTLFCFLFAFSSVLDVTLRPKEKNVVEYSETQELLRTIEINPLRIYGVVCATKIMKLR 656
 DB 608 DSTLFCFLFAFSSVLDVTLRPKEKNVVEYSETQELLRTIEINPLRIYGVVCATKIMKLR 667
 QY 657 KILEKVEAASGFTSEKPEEFLNLFHILRLVEPLLKIRSAQKQVQDCYFYQIFMEKNE 716
 DB 668 KILEKVEAASGFTSEKPEEFLNLFHILRLVEPLLKIRSAQKQVQDCYFYQIFMEKNE 727
 QY 717 KVGVTPIQOLLEWSFNSLNKFAEAPSCLIQMPREGKDFKLPKIPPSLELNTIDLLED 776
 DB 728 KVGVTPIQOLLEWSFNSLNKFAEAPSCLIQMPREGKDFKLPKIPPSLELNTIDLLED 787

QY 777 TPQCRICGGLAMYECECYDDPDISAGKIQKCTCNTQVHLHPKRLNHNKYNPVSXPKD 836
 DB 788 TPQCRICGGLAMYECECYDDPDISAGKIQKCTCNTQVHLHPKRLNHNKYNPVSXPKD 847
 QY 837 LPDWDWRHGCIPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLFFDSDMADRGGQNGFNI 896
 DB 848 LPDWDWRHGCIPQNNMELFAVLCTIETSHYVAFVKYKDDSAWLFFDSDMADRGGQNGFNI 907
 QY 897 PQVTPCEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 949
 DB 908 PQVTPCEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 960

RESULT 7
 PCT-US02-14570-4
 ; Sequence 4, Application PC/TUS0214570
 ; GENERAL INFORMATION:
 ; APPLICANT: IMMUNEX CORPORATION
 ; APPLICANT: Derry, Jonathan M. J.
 ; APPLICANT: Fanslow III, William
 ; APPLICANT: Dougall, William C.
 ; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
 ; FILE REFERENCE: 3198-WO
 ; CURRENT APPLICATION NUMBER: PCT/US02/14570
 ; CURRENT FILING DATE: 2002-05-07
 ; PRIOR APPLICATION NUMBER: 09/851,673
 ; PRIOR FILING DATE: 2001-05-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 956
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-14570-4

Query Match 98.8%; Score 4971.5; DB 1; Length 956;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

QY 1 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 60
 DB 1 MSSGLWSQEKVTSFYWEERIFVLLQECSTVDKQTKLLKVPKSGSIGQYIQDRSVGHSRI 60
 QY 61 PSAKKNQIGLKILEQPHAVLFVDE-DVVEINEKFTTELLAITNCEERFSLFKNRNLS 119
 DB 61 PSAKKNQIGLKILEQPHAVLFVDEKDVVEINEKFTTELLAITNCEERFSLFKNRNLS 120
 QY 120 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 179
 DB 121 KGLQIDVGCPCVKQLRSGBEKEFGVVRFRGPLLAERTVSGIFPGVELLEGRGQGTGTV 180
 QY 180 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPLEINSRVSXKG 238
 DB 181 YQKQLFQDCEDCG-FVALDKLEIEDDDTALESYAGPGDTMQLPPLPLEINSRVSXKG 240
 QY 239 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHN 297
 DB 241 GTTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHN 300
 QY 298 DIIPESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNR-RSELPFTYTLNG 353
 DB 301 DIIPESVTOERRPPKLAFAFMRGVDKSGSSHNKPKATGSTSDPGNR-RSELPFTYTLNG 360
 QY 354 SSVDSQPSKSNVTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSLP 413
 DB 361 SSVDSQPSKSNVTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSLTNENRPHSLP 420
 QY 414 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMP-PGNSHGLEVGSIAEVKE 473
 DB 421 FSLTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPPPLAMP-PGNSHGLEVGSIAEVKE 480
 QY 474 NPPFYGVIRWIOQPPGLNEVLVAGLEDEBCAGCTDGTFRGTRYFTCALKKALFVKLKSCR 533

Db 481 NPPFYGIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 540
QY 534 PDSRFASLPQVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 593
Db 541 PDSRFASLPQVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 600
QY 594 CYLDSTFLCLFAFSSVLDTVLLRPKEKNDVEYVSETOELLRTIENVPLRIYGVVCATKIM 553
Db 601 CYLDSTFLCLFAFSSVLDTVLLRPKEKNDVEYVSETOELLRTIENVPLRIYGVVCATKIM 660
QY 654 KLRKILEKVEAASGFTSEEDPEEFNLILFHILRVEPLLKRSAGOKVQDCVFIQIFME 713
Db 661 KLRKILEKVEAASGFTSEEDPEEFNLILFHILRVEPLLKRSAGOKVQDCVFIQIFME 720
QY 714 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 773
Db 721 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 780
QY 774 LEDTPRQCRICGGLAMVRECYDDPDISAGKIQKQCTCNTQVHLHPKRLNHNKYNPVS 833
Db 781 LEDTPRQCRICGGLAMVRECYDDPDISAGKIQKQCTCNTQVHLHPKRLNHNKYNPVS 840
QY 834 PKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGQNG 893
Db 841 PKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGQNG 900
QY 894 FNIQVTPCPEVGEYLKMSLEDLHSLDSRRIOQCARRLLCDAYMCYQSTMSLYK 949
Db 901 FNIQVTPCPEVGEYLKMSLEDLHSLDSRRIOQCARRLLCDAYMCYQSTMSLYK 956

RESULT 8

US-09-851-673-4
; Sequence 4, Application US/09851673
; GENERAL INFORMATION:
; APPLICANT: Derry, Jonathan
; APPLICANT: Fanslow, William
; APPLICANT: Dougall, William
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
; FILE REFERENCE: 3198
; CURRENT APPLICATION NUMBER: US/09/851,673 } late
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-673-4

Query Match 98.8%; Score 4971.5; DB 23; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;
QY 1 MSSGLWSQEKVTSPPYWEERIFYLLLOECSTVDKQTKLLKVPKGSIGQYIQDRSVGHSRI 60
Db 1 MSSGLWSQEKVTSPPYWEERIFYLLLOECSTVDKQTKLLKVPKGSIGQYIQDRSVGHSRI 60
QY 61 PSAGKKNQIGLKILQPHAVLPVDE-DVVEINEKTELLLAITNCEERFSLPKNRLS 119
Db 61 PSAGKKNQIGLKILQPHAVLPVDEKDVVEINEKTELLLAITNCEERFSLPKNRLS 120
QY 120 KGLQIDVGPVKVQLSGEKEFGVVRFGPPLAERTVSGIFPGVLLBEGRGQGTGV 179
Db 121 KGLQIDVGPVKVQLSGEKEFGVVRFGPPLAERTVSGIFPGVLLBEGRGQGTGV 180
QY 180 YQKQLFQCDDECG-FVALDKLEIEDDTALESYAGTDMQVLPPLPINSRVSLKG 238
Db 181 YQKQLFQCDDECGFVALDKLEIEDDTALESYAGTDMQVLPPLPINSRVSLKV 240
QY 239 GETIESGTWIFCDVLPKGSLSGYFVGVDMDNPNIGNWGRPDGV-LCSFACVESTILLHN 297
Db 241 GETIESGTWIFCDVLPKGSLSGYFVGVDMDNPNIGNWGRPDGV-LCSFACVESTILLHN 300

QY 298 DIIP---ESVTOERRPPKLAFMRSRGVGDKGSSSHNKPATGSTDGPNR-RSELYFTLNG 353
Db 301 DIIPALSESVTQERRPPKLAFMRSRGVGDKGSSSHNKPATGSTDGPNRNRSELYFTLNG 360
QY 354 SSVDSOQSKSNKWTYIDEVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTTERPHSLP 413
Db 361 SSVDSOQSKSNKWTYIDEVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTTERPHSLP 420
QY 414 FSLTKMPNTNGSIHGSPGLSPLSAQSVMEELNTAPQESPPPLAMPNGSHGLEVSLEAVKE 473
Db 421 FSLTKMPNTNGSIHGSPGLSPLSAQSVMEELNTAPQESPPPLAMPNGSHGLEVSLEAVKE 480
QY 474 NPPFYGIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 533
Db 481 NPPFYGIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKKALFVKLKSCR 540
QY 534 PDSRFASLPQVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 593
Db 541 PDSRFASLPQVNSQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNS 600
QY 594 CYLDSTFLCLFAFSSVLDTVLLRPKEKNDVEYVSETOELLRTIENVPLRIYGVVCATKIM 653
Db 601 CYLDSTFLCLFAFSSVLDTVLLRPKEKNDVEYVSETOELLRTIENVPLRIYGVVCATKIM 660
QY 654 KLRKILEKVEAASGFTSEEDPEEFNLILFHILRVEPLLKRSAGOKVQDCVFIQIFME 713
Db 661 KLRKILEKVEAASGFTSEEDPEEFNLILFHILRVEPLLKRSAGOKVQDCVFIQIFME 720
QY 714 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 773
Db 721 KNEKVGVPITQQLLEWSFINSNLKFAEAPSCIIQMPRFKGFVKLKKIPPSLELNITDL 780
QY 774 LEDTPRQCRICGGLAMVRECYDDPDISAGKIQKQCTCNTQVHLHPKRLNHNKYNPVS 833
Db 781 LEDTPRQCRICGGLAMVRECYDDPDISAGKIQKQCTCNTQVHLHPKRLNHNKYNPVS 840
QY 834 PKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGQNG 893
Db 841 PKDLPDMDWRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLPFDSMADRDGQNG 900
QY 894 FNIQVTPCPEVGEYLKMSLEDLHSLDSRRIOQCARRLLCDAYMCYQSTMSLYK 949
Db 901 FNIQVTPCPEVGEYLKMSLEDLHSLDSRRIOQCARRLLCDAYMCYQSTMSLYK 956

RESULT 9

US-10-755-889-490
; Sequence 490, Application US/10755889
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 490
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-490

Query Match 98.8%; Score 4971.5; DB 33; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;
QY 1 MSSGLWSQEKVTSPPYWEERIFYLLLOECSTVDKQTKLLKVPKGSIGQYIQDRSVGHSRI 60

```
Db 1 MSSGLWSQEKVTSFYWEERIFYLLOQCSVTDKQTKLLKVPKSGISQYIQRDSVGHRI 60
QY 61 PSAKGKKNQIGLKILEQPHAVLFVDE-DVVEINEKFTTELLAITNCEERFSLFKNRRLS 119
Db 61 PSAKGKKNQIGLKILEQPHAVLFVDEKDVVEINEKFTTELLAITNCEERFSLFKNRRLS 120
QY 120 KGLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQGTGTV 179
Db 121 KGLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQGTGTV 180
QY 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDMQVLPLEINSRVS LKG 238
Db 181 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDMQVLPLEINSRVS LKV 240
QY 239 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 297
Db 241 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 300
QY 298 DIIP---ESVTOERRPPKLAFAFMRSGVGDGSSSHNPKATGSTSDPGNR-RSELFYTLNG 353
Db 301 DIIPALSESVTOERRPPKLAFAFMRSGVGDGSSSHNPKATGSTSDPGNRNRSSELYTLNG 360
QY 354 SSVDSQPSQSKNNTWYIDVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTENRHS LIP 413
Db 361 SSVDSQPSQSKNNTWYIDVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTENRHS LIP 420
QY 414 FSLTKMPNTNGSIHSPSLSAQSVMELNTPAQVSPPLAMPNGSHGLEVGS LAE VKE 473
Db 421 FSLTKMPNTNGSIHSPSLSAQSVMELNTPAQVSPPLAMPNGSHGLEVGS LAE VKE 480
QY 474 NPPFYGVIRWIGOPPGNLVLALEDEACAGCTDGTFRGTRFTCALKKALFVKL KSCR 533
Db 481 NPPFYGVIRWIGOPPGNLVLALEDEACAGCTDGTFRGTRFTCALKKALFVKL KSCR 540
QY 534 PDSRFASLQPVSNQIERCNSLAFGGYLSVEVENTPPKMEKEGLEIMIGKKGIQGHYNS 593
Db 541 PDSRFASLQPVSNQIERCNSLAFGGYLSVEVENTPPKMEKEGLEIMIGKKGIQGHYNS 600
QY 594 CYLDSTLFCFLAFSSVLDTVLLRPKEKNVVEYSETQELLRTEIVNPLRIYGVCA TKIM 653
Db 601 CYLDSTLFCFLAFSSVLDTVLLRPKEKNVVEYSETQELLRTEIVNPLRIYGVCA TKIM 660
QY 654 KLRKILEKVEAASGFTSEKDPPEEFNLIFHHILRVEPLLKIRSAQKQVDCYFYQI FME 713
Db 661 KLRKILEKVEAASGFTSEKDPPEEFNLIFHHILRVEPLLKIRSAQKQVDCYFYQI FME 720
QY 714 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKLPKKIPPSLEL NITDL 773
Db 721 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKLPKKIPPSLEL NITDL 780
QY 774 LEDTPROCRICGGLAMYECCRECYDDPDISAGKIQKQCKTQNTQVHLHPKRLNHNKYNP VSL 833
Db 781 LEDTPROCRICGGLAMYECCRECYDDPDISAGKIQKQCKTQNTQVHLHPKRLNHNKYNP VSL 840
QY 834 PKDLDPMDWRHGCIPQNNMELFAVLCTIETSHYVAFVKYKQKDSAWLFFDSDMADR DGGQNG 893
Db 841 PKDLDPMDWRHGCIPQNNMELFAVLCTIETSHYVAFVKYKQKDSAWLFFDSDMADR DGGQNG 900
QY 894 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLICDAYCMYQSTWSLYK 949
Db 901 FNIPOVTPCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLICDAYCMYQSTWSLYK 956
```

RESULT 10

US-60-440-068-490

```
; Sequence 490, Application US/60440068
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CARMAN, JULIE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE
; TITLE OF INVENTION: NF-KB PATHWAY
; FILE REFERENCE: 3053-4191
```

```
; CURRENT APPLICATION NUMBER: US/60/440,068
; CURRENT FILING DATE: 2003-01-14
```

; NUMBER OF SEQ ID NOS: 746

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 490

; LENGTH: 956

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-440-068-490

Query Match 98.8%; Score 4971.5; DB 37; Length 956;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;

```
QY 1 MSSGLWSQEKVTSFYWEERIFYLLOQCSVTDKQTKLLKVPKSGISQYIQRDSVGHRI 60
Db 1 MSSGLWSQEKVTSFYWEERIFYLLOQCSVTDKQTKLLKVPKSGISQYIQRDSVGHRI 60
QY 61 PSAKGKKNQIGLKILEQPHAVLFVDE-DVVEINEKFTTELLAITNCEERFSLFKNRRLS 119
Db 61 PSAKGKKNQIGLKILEQPHAVLFVDEKDVVEINEKFTTELLAITNCEERFSLFKNRRLS 120
QY 120 KGLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQGTGTV 179
Db 121 KGLQIDVGCPCVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVLLBEGRGQGTGTV 180
QY 180 YQKQLFQCDDECG-FVALDKLELIEDDDTALSDYAGPDMQVLPLEINSRVS LKG 238
Db 181 YQKQLFQCDDECGFVALDKLELIEDDDTALSDYAGPDMQVLPLEINSRVS LKV 240
QY 239 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 297
Db 241 GETIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 300
QY 298 DIIP---ESVTOERRPPKLAFAFMRSGVGDGSSSHNPKATGSTSDPGNR-RSELFYTLNG 353
Db 301 DIIPALSESVTOERRPPKLAFAFMRSGVGDGSSSHNPKATGSTSDPGNRNRSSELYTLNG 360
QY 354 SSVDSQPSQSKNNTWYIDVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTENRHS LIP 413
Db 361 SSVDSQPSQSKNNTWYIDVAEDPAKSLTEISTDFDRSSPPLOPPVNSLTENRHS LIP 420
QY 414 FSLTKMPNTNGSIHSPSLSAQSVMELNTPAQVSPPLAMPNGSHGLEVGS LAE VKE 473
Db 421 FSLTKMPNTNGSIHSPSLSAQSVMELNTPAQVSPPLAMPNGSHGLEVGS LAE VKE 480
QY 474 NPPFYGVIRWIGOPPGNLVLALEDEACAGCTDGTFRGTRFTCALKKALFVKL KSCR 533
Db 481 NPPFYGVIRWIGOPPGNLVLALEDEACAGCTDGTFRGTRFTCALKKALFVKL KSCR 540
QY 534 PDSRFASLQPVSNQIERCNSLAFGGYLSVEVENTPPKMEKEGLEIMIGKKGIQGHYNS 593
Db 541 PDSRFASLQPVSNQIERCNSLAFGGYLSVEVENTPPKMEKEGLEIMIGKKGIQGHYNS 600
QY 594 CYLDSTLFCFLAFSSVLDTVLLRPKEKNVVEYSETQELLRTEIVNPLRIYGVCA TKIM 653
Db 601 CYLDSTLFCFLAFSSVLDTVLLRPKEKNVVEYSETQELLRTEIVNPLRIYGVCA TKIM 660
QY 654 KLRKILEKVEAASGFTSEKDPPEEFNLIFHHILRVEPLLKIRSAQKQVDCYFYQI FME 713
Db 661 KLRKILEKVEAASGFTSEKDPPEEFNLIFHHILRVEPLLKIRSAQKQVDCYFYQI FME 720
QY 714 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKLPKKIPPSLEL NITDL 773
Db 721 KNEKVGVPITIQOLLEWSFINSNLKFAEAPSLIIQMPRFGKDFKLPKKIPPSLEL NITDL 780
QY 774 LEDTPROCRICGGLAMYECCRECYDDPDISAGKIQKQCKTQNTQVHLHPKRLNHNKYNP VSL 833
Db 781 LEDTPROCRICGGLAMYECCRECYDDPDISAGKIQKQCKTQNTQVHLHPKRLNHNKYNP VSL 840
QY 834 PKDLDPMDWRHGCIPQNNMELFAVLCTIETSHYVAFVKYKQKDSAWLFFDSDMADR DGGQNG 893
Db 841 PKDLDPMDWRHGCIPQNNMELFAVLCTIETSHYVAFVKYKQKDSAWLFFDSDMADR DGGQNG 900
```


QY 894 FNIPOVTPCPVEGYLKMSEDLHLSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 949
Db 901 FNIPOVTPCPVEGYLKMSEDLHLSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 956
RESULT 11
US-60-469-757-490
; Sequence 490, Application US/60469757
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE REFERENCE: D0284 P5P1
; CURRENT APPLICATION NUMBER: US/60/469,757
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 490
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-469-757-490
Query Match 98.8%; Score 4971.5; DB 37; Length 956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 948; Conservative 0; Mismatches 1; Indels 7; Gaps 5;
QY 1 MSSGLWSQKVTSPWEEIFVLLLOECSTVTKQTKLLKVPKSGISQYIQRSVGHSHI 60
Db 1 MSSGLWSQKVTSPWEEIFVLLLOECSTVTKQTKLLKVPKSGISQYIQRSVGHSHI 60
QY 61 PSAKGKNOIGLKILOPHAVLFVDE-DVVEINEKFTELLAITNCEERFSLFKNNRNL 119
Db 61 PSAKGKNOIGLKILOPHAVLFVDEKDVVEINEKFTELLAITNCEERFSLFKNNRNL 120
QY 120 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIPFGVELLEBEGRGQFTDG 179
Db 121 KGLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIPFGVELLEBEGRGQFTDG 180
QY 180 YQKQLFOCDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLSK 238
Db 181 YQKQLFOCDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLSK 240
QY 239 GETIESGTVIFCDVLPKESLSGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 297
Db 241 GETIESGTVIFCDVLPKESLSGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 300
QY 298 DIIP-----ESVTQERRPPKLAFMRSRGVDKGSSSHNPKATGSTSDPCNR-RSELFYTLNG 353
Db 301 DIIPALSESVTQERRPPKLAFMRSRGVDKGSSSHNPKATGSTSDPCNRNRSFLYTLNG 360
QY 354 SSVDSQPSKSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRFHSPL 413
Db 361 SSVDSQPSKSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRFHSPL 420
QY 414 FSLTKMNTNGSCHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSIAEVE 473
Db 421 FSLTKMNTNGSCHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSIAEVE 480
QY 474 NPPFYGVIRWIGPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKLSKCR 533
Db 481 NPPFYGVIRWIGPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKLSKCR 540
QY 534 PDSRFASLQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKGIQGHYNS 593
Db 541 PDSRFASLQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKGIQGHYNS 600
QY 594 CYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYSETQELLRTETVNPRIYGVVCATKIM 653
Db 601 CYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYSETQELLRTETVNPRIYGVVCATKIM 660
QY 654 KLRKILEKVEAASGFTSEKDPPEFLNLPFHILRVEPLLKIRSAQKVQDCFYQIFME 713

Db 661 KURKILEKVEAASGFTSEKDPPEFLNLPFHILRVEPLLKIRSAQKVQDCFYQIFME 720
QY 714 KNEKVGVPITIQOLLEWSFINSLKFAEAPSCLIIOMPRFGKQFKLKKIIPPSLELNITDL 773
Db 721 KNEKVGVPITIQOLLEWSFINSLKFAEAPSCLIIOMPRFGKQFKLKKIIPPSLELNITDL 780
QY 774 LEDTPRQCRICGGLAMYECCREYDDPDISAGKIKQFCCKTQNTQVHLHPKLNHNKYNPVS 833
Db 781 LEDTPRQCRICGGLAMYECCREYDDPDISAGKIKQFCCKTQNTQVHLHPKLNHNKYNPVS 840
QY 834 PKDLDPMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLFFDSMARDGGONG 893
Db 841 PKDLDPMDWRHGCIPCONMELFAVLCTIETSHYVAFVKYKDDSAWLFFDSMARDGGONG 900
QY 894 FNIPOVTPCPVEGYLKMSEDLHLSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 949
Db 901 FNIPOVTPCPVEGYLKMSEDLHLSLDSRRIOGCARRLLCDAYMCYQSTMSLYK 956
RESULT 12
PCT-US01-01239-1743
; Sequence 1743, Application PC/TUS0101239
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ13PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01239
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1743
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01239-1743
Query Match 77.0%; Score 3876.5; DB 1; Length 739;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 33; Gaps 4;
QY 184 QLFQCDDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLSKGETI 242
Db 1 QLFQCDDEDCG-FVALDKLEIEDDDTALESYAGPGDTMQVLPPLPINSRVSLSK 242
QY 243 ESGTVIFCDVLPKESLSGYFVGVDMDNPIGNWDGRFDGV-LCSFACVESTILLHIN 301
Db 56 -----DNPIGNWDGRFDGV-LCSFACVESTILLHIN 301
QY 302 ESVTQERRPPKLAFMRSRGVDKGSSSHNPKATGSTSDPCNR-RSELFYTLNGSSVDSQP 360
Db 91 ESVTQERRPPKLAFMRSRGVDKGSSSHNPKATGSTSDPCNRNRSFLYTLNGSSVDSQP 150
QY 361 QSKSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRFHSPLTKMP 420
Db 151 QSKSKNTWIDEVAEDPAKSLTEISTDFRSPPLPQPPVNSLTNRRFHSPLTKMP 210
QY 421 NTNGSIGHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSIAEVEKPNPPFYGV 480
Db 211 NTNGSIGHSPSLSAQSVMEELNTAPVQESPPPLAMPNGSHGLEVGSIAEVEKPNPPFYGV 270
QY 481 IRWIGOPPGNEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKLSKCRDPSRPAS 540
Db 271 IRWIGOPPGNEVLAGELEDECACTDGTFRGTRYFTCALKKALFVKLSKCRDPSRPAS 330
QY 541 LQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600
Db 331 LQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 390
QY 601 FCLFAFSSVLDTVLLRPKEKNDVEYSETQELLRTETVNPRIYGVVCATKIMKRLKILE 660
Db 391 FCLFAFSSVLDTVLLRPKEKNDVEYSETQELLRTETVNPRIYGVVCATKIMKRLKILE 450

QY 661 KVEAASGFTSEEDKPEEFNLHILRVEPLKIRSAQKQVDCYFYQIFMEKNEKVG 720
Db 451 KVEAASGFTSEEDKPEEFNLHILRVEPLKIRSAQKQVDCYFYQIFMEKNEKVG 510
QY 721 PTIQOLLEWSFNSNLKFAEAPSCIIQMPRFGKDFKFKKIPPSLELNITDLEDTPRQ 780
Db 511 PTIQOLLEWSFNSNLKFAEAPSCIIQMPRFGKDFKFKKIPPSLELNITDLEDTPRQ 570
QY 781 CRICGGLAMWECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 840
Db 571 CRICGGLAMWECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 630
QY 841 DWRHGCIPQNMELFAVLCIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 900
Db 631 DWRHGCIPQNMELFAVLCIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 690
QY 901 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYCMYQSPMTSLYK 949
Db 691 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYCMYQSPMTSLYK 739

RESULT 13
US-09-764-902-1743
; Sequence 1743, Application US/09764902
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213
; CURRENT APPLICATION NUMBER: US/09/764,902
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1743
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-902-1743

Query Match 77.0%; Score 3876.5; DB 22; Length 739;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 33; Gaps 4;

QY 184 QLFQCDDECG-FVALDKLEIEDDDTALSDYAGPGDTMQVELPPLPINSRSLKGGGTI 242
Db 1 QLFQCDDECGFVALDKLEIEDDDTALSDYAGPGDTMQVELPPLPINSRSLKGGGTI 55
QY 243 ESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRPDGV-LCSFACVESTILLHINDIIP 301
Db 56 -----DNPICNWDGRPDGVQLCSFACVESTILLHINDIIP 90
QY 302 ESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSDPQNR-RSELFTLANGSSVDSOP 360
Db 91 ESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSDPQNRNRSSELFTLANGSSVDSOP 150
QY 361 QSKSKNTWYIDEAEDPAKSLTEISDFDRSPPLQPPVNSLTNTRFHSLPFSITKMP 420
Db 151 QSKSKNTWYIDEAEDPAKSLTEISDFDRSPPLQPPVNSLTNTRFHSLPFSITKMP 210
QY 421 NTNGSIGHSPLSLSAQSVMEELNTPVQSPPLAMPNGSHGLEVGSIAEVKENPPFYGV 480
Db 211 NTNGSIGHSPLSLSAQSVMEELNTPVQSPPLAMPNGSHGLEVGSIAEVKENPPFYGV 270
QY 481 IRWIQOPGLNVLAVGLEDEACAGTDTGTFRTGTRFTCALKALPVKLKSCRPSRFPAS 540
Db 271 IRWIQOPGLNVLAVGLEDEACAGTDTGTFRTGTRFTCALKALPVKLKSCRPSRFPAS 330
QY 541 LQPVSNQIERCNLSAFGGVLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 600
Db 331 LQPVSNQIERCNLSAFGGVLSVVEENTPPKMEKEGLEIMIGKKGIQGHYNSCYLDSTL 390
QY 601 FCLFAFSSVLDTVLLRPKEKNDVEYSETQELLRTIIVNPLRIYGVVCAIKMKRILE 660

Db 391 FCLFAFSSVLDTVLLRPKEKNDVEYSETQELLRTIIVNPLRIYGVVCAIKMKRILE 450
QY 661 KVEAASGFTSEEDKPEEFNLHILRVEPLKIRSAQKQVDCYFYQIFMEKNEKVG 720
Db 451 KVEAASGFTSEEDKPEEFNLHILRVEPLKIRSAQKQVDCYFYQIFMEKNEKVG 510
QY 721 PTIQOLLEWSFNSNLKFAEAPSCIIQMPRFGKDFKFKKIPPSLELNITDLEDTPRQ 780
Db 511 PTIQOLLEWSFNSNLKFAEAPSCIIQMPRFGKDFKFKKIPPSLELNITDLEDTPRQ 570
QY 781 CRICGGLAMWECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 840
Db 571 CRICGGLAMWECRECYDDPDISAGKIKQFCKTCTQVHLHPKRLNHNKYNPVSPLPKDLPDW 630
QY 841 DWRHGCIPQNMELFAVLCIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 900
Db 631 DWRHGCIPQNMELFAVLCIETSHYVAFVYKGDSDSAWLPFDSMADRDGGQNGFNIPQVT 690
QY 901 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYCMYQSPMTSLYK 949
Db 691 PCPEVGEYLKMSLEDLHSLDSRRIOGCARRLLCDAYCMYQSPMTSLYK 739

RESULT 14
US-09-786-797B-9
; Sequence 9, Application US/09786797B
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/09/786,797B
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2363327
US-09-786-797B-9

Query Match 76.7%; Score 3862; DB 22; Length 731;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 221 MQVELPPLPINSRSLKGGGTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRPDG 280
Db 1 MQVELPPLPINSRSLKGGGTIESGTVIFCDVLPKESLGYFVGVDMDNPIGNWDRPDG 60
QY 281 V-LCSFACVESTILLHINDIIPESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSD 339
Db 61 VQLCSFACVESTILLHINDIIPESVTOERRPPKLAFMRSRGVGDGSSSHNPKATGSTSD 120
QY 340 PGNR-RSELFTLANGSSVDSOPQSKNTWYIDEAEDPAKSLTEISTDFDRSSPPPLQPP 398
Db 121 PGNRNSLFTLANGSSVDSOPQSKNTWYIDEAEDPAKSLTEISTDFDRSSPPPLQPP 180
QY 399 PVNSLTNTRFHSLPFSITKMPNTNGSIGHSPLSLSAQSVMEELNTPVQSPPLAMP 458

```
Db 181 PVNSLTENRHSLSLPSLTKMNTNGSIHSPSLSSAQSVMEELNTPVQESPPPLAMPFG 240
Qy 459 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTGTRGTYFT 518
Db 241 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTGTRGTYFT 300
Qy 519 CALKKALFVKLKSRCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 578
Db 301 CALKKALFVKLKSRCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 360
Qy 579 IMIGKKGIQGHYNSCYLDSTLFCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 638
Db 361 IMIGKKGIQGHYNSCYLDSTLFCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 420
Qy 639 NPLRIYGVYCATKIMKRLKILEKVAASGFTSEEDPEEFNLIFHILRVEPLLKIRSA 698
Db 421 NPLRIYGVYCATKIMKRLKILEKVAASGFTSEEDPEEFNLIFHILRVEPLLKIRSA 480
Qy 699 GOKVQDCYFYQIFMEKNEKVGVTIQQLLWFSFINSNLKFAEAPSCLI IQMPRFGKDFKL 758
Db 481 GOKVQDCYFYQIFMEKNEKVGVTIQQLLWFSFINSNLKFAEAPSCLI IQMPRFGKDFKL 540
Qy 759 FKIPFSLNLTDLLEDTTPQCRI CGGLAMYECRECYDDPDISAGKIKQFCKTCTQVH 818
Db 541 FKIPFSLNLTDLLEDTTPQCRI CGGLAMYECRECYDDPDISAGKIKQFCKTCTQVH 600
Qy 819 LHPKRLNHNKYNPVS LKPLDWDWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAW 878
Db 601 LHPKRLNHNKYNPVS LKPLDWDWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAW 660
Qy 879 LFFDSMADRDGGQNGFNIPQVTPCPEVGEYLYKMSLEDLHSLDSRR IQGCARRLLCDAYMC 938
Db 661 LFFDSMADRDGGQNGFNIPQVTPCPEVGEYLYKMSLEDLHSLDSRR IQGCARRLLCDAYMC 720
Qy 939 MYQSPMTSLYK 949
Db 721 MYQSPMTSLYK 731
```

RESULT 15

```
US-10-921-707-9
; Sequence 9, Application US/10921707
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/10/921,707
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: US/09/786,797
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2363327
US-10-921-707-9
```

```
Query Match 76.7%; Score 3862; DB 35; Length 731;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 221 MVELPPLPINSRVS LKGGETIESGTVIFCDVLP GKESLG YFVGVDMDNPIGNWDGDFDG 280
Db 1 MVELPPLPINSRVS LKGGETIESGTVIFCDVLP GKESLG YFVGVDMDNPIGNWDGDFDG 60
Qy 281 V-LCSFACVESITILLHINDIIIPESVTQERRPPKLPFMSRGVGDGKSSSHNKPATGTS 339
Db 61 VQ-LCSFACVESITILLHINDIIIPESVTQERRPPKLPFMSRGVGDGKSSSHNKPATGTS 120
Qy 340 PGNR-RSELFYTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSPPIQPP 398
Db 121 PGNR-SELFTYTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSPPIQPP 180
Qy 399 PVNSLTENRHSLSLPSLTKMNTNGSIHSPSLSSAQSVMEELNTPVQESPPPLAMPFG 458
Db 181 PVNSLTENRHSLSLPSLTKMNTNGSIHSPSLSSAQSVMEELNTPVQESPPPLAMPFG 240
Qy 459 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTGTRGTYFT 518
Db 241 NSHGLEVSLAEVKNPPYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTGTRGTYFT 300
Qy 519 CALKKALFVKLKSRCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 578
Db 301 CALKKALFVKLKSRCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKEGLE 360
Qy 579 IMIGKKGIQGHYNSCYLDSTLFCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 638
Db 361 IMIGKKGIQGHYNSCYLDSTLFCFLAFSSVLDTVLLRPKEKNDVEYYSETQELLRTETV 420
Qy 639 NPLRIYGVYCATKIMKRLKILEKVAASGFTSEEDPEEFNLIFHILRVEPLLKIRSA 698
Db 421 NPLRIYGVYCATKIMKRLKILEKVAASGFTSEEDPEEFNLIFHILRVEPLLKIRSA 480
Qy 699 GOKVQDCYFYQIFMEKNEKVGVTIQQLLWFSFINSNLKFAEAPSCLI IQMPRFGKDFKL 758
Db 481 GOKVQDCYFYQIFMEKNEKVGVTIQQLLWFSFINSNLKFAEAPSCLI IQMPRFGKDFKL 540
Qy 759 FKIPFSLNLTDLLEDTTPQCRI CGGLAMYECRECYDDPDISAGKIKQFCKTCTQVH 818
Db 541 FKIPFSLNLTDLLEDTTPQCRI CGGLAMYECRECYDDPDISAGKIKQFCKTCTQVH 600
Qy 819 LHPKRLNHNKYNPVS LKPLDWDWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAW 878
Db 601 LHPKRLNHNKYNPVS LKPLDWDWRHGCIPQCNMELFAVLCTIETSHYVAFVYKGDSDSAW 660
Qy 879 LFFDSMADRDGGQNGFNIPQVTPCPEVGEYLYKMSLEDLHSLDSRR IQGCARRLLCDAYMC 938
Db 661 LFFDSMADRDGGQNGFNIPQVTPCPEVGEYLYKMSLEDLHSLDSRR IQGCARRLLCDAYMC 720
Qy 939 MYQSPMTSLYK 949
Db 721 MYQSPMTSLYK 731
```

Search completed: April 18, 2005, 15:32:41
Job time : 359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2005, 15:35:41 ; Search time 7161 Seconds
(without alignments)
5410.008 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 5034

Sequence: 1 MSSGLWSQKVTSPWBERI.....RLLCDAYMCVQSPWLSLYK 949

Scoring table:

Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09671687/runat_18042005_115019_17191/app_query.fasta_1.1095
-DB=Pending_Patents_NA_Main_QPMT=fastap -SUFFIX=rnmp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09671687@cgn 1 1 6376 @runat_18042005_115019_17191 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US08_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US09_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*

29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US100C_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US105A_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US105B_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US106A_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US107A_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US107B_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US107C_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US107D_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US108A_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US108B_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US109A_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US109B_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US109C_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US110_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*

102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
 103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
 104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
 105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
 106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
 107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
 108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
 109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
 110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
 111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
 112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
 113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
 114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
 115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
 116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
 117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4992	99.2	3715	30	US-09-671-687A-2
2	4983	99.0	4661	23	US-09-522-303-1293
3	4983	99.0	4664	20	US-09-315-788-1729
4	4983	99.0	4664	20	US-09-315-788-1729
5	4983	99.0	4664	34	US-09-783-514-1729
6	4983	99.0	4664	34	US-09-783-514A-1729
7	4983	99.0	4668	21	US-09-396-087-4309
8	4983	99.0	4668	21	US-09-396-970-7559
9	4983	99.0	4668	21	US-09-397-424-5019
10	4983	99.0	4668	21	US-09-397-424A-5019
11	4983	99.0	4668	22	US-09-432-241A-3760
12	4983	99.0	4668	22	US-09-434-737-1278
13	4983	99.0	4668	36	US-09-850-118-1278
14	4983	99.0	6831	1	PCT-US02-27777-49
15	4983	99.0	6831	1	PCT-US02-27777A-49
16	4983	99.0	6871	49	US-10-170-235-39761
17	4971.5	98.8	5371	36	US-09-851-673-3
18	4971.5	98.8	5371	36	US-10-755-889-489
19	4971.5	98.8	5371	62	US-10-887-553A-199
20	4971.5	98.8	5371	64	US-10-887-553A-199
21	4971.5	98.8	5371	114	US-60-440-068-489
22	4971.5	98.8	5371	116	US-60-469-757-489
23	4970	98.7	6314	102	US-60-324-185-23812
24	4953	98.4	6315	90	US-60-213-359-5981
25	4953	98.4	6316	97	US-60-278-232-7384
26	4942	98.2	4527	34	US-09-786-797B-25
27	4942	98.2	4527	65	US-10-921-707-25
28	4942	98.2	4527	82	US-60-131-321-21
29	4916	97.7	6315	86	US-60-172-360-22693
30	4601	91.4	2845	28	US-09-629-469A-18842
31	4601	91.4	2845	65	US-10-917-503-18842
32	3876.5	77.0	4716	1	PCT-US01-01239-843
33	3876.5	77.0	4716	34	US-09-764-902-843
34	3676	73.0	2523	22	US-09-488-725A-613
35	3676	73.0	2523	46	US-10-037-270-290

36 3676 73.0 2523 48 US-10-105-837-290
 37 3676 73.0 2523 48 US-10-117-722-290
 38 3676 73.0 2523 51 US-10-258-898A-613
 39 3676 73.0 2523 51 US-10-286-897-613
 40 3526 70.0 4493 48 US-10-144-771-8151
 41 3526 70.0 4493 106 US-60-360-207-8151
 42 3523 70.0 4286 22 US-09-471-275-2028
 43 3523 70.0 4286 22 US-09-488-725A-4185
 44 3523 70.0 4286 22 US-09-488-725B-6145
 45 3523 70.0 4286 26 US-09-552-317-6145

ALIGNMENTS

RESULT 1

US-09-671-687A-2
 ; Sequence 2, Application US/09671687A
 ; GENERAL INFORMATION:
 ; APPLICANT: WALLACH, David
 ; APPLICANT: KOVALENKO, Andrei
 ; APPLICANT: CANTARELLA, Giuseppeina
 ; TITLE OF INVENTION: INHIBITOR OF NF-KB ACTIVATOR
 ; FILE REFERENCE: WALLACH-25
 ; CURRENT APPLICATION NUMBER: US/09/671,687A
 ; CURRENT FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: 09/646,403
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: IL 126024
 ; PRIOR FILING DATE: 1998-09-01
 ; PRIOR APPLICATION NUMBER: IL 134604
 ; PRIOR FILING DATE: 2000-02-17
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 3715
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (18)..(18)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; NAME/KEY: misc feature
 ; LOCATION: (22)..(22)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; NAME/KEY: misc feature
 ; LOCATION: (756)..(756)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; NAME/KEY: misc feature
 ; LOCATION: (1348)..(1348)
 ; OTHER INFORMATION: n is either a, c, g, or t.
 ; US-09-671-687A-2

Alignment Scores:
 Pred. No.: 0 Length: 3715
 Score: 4992.00 Matches: 949
 Percent Similarity: 99.58% Conservativeness: 0
 Best Local Similarity: 99.58% Mismatches: 0
 Query Match: 99.17% Indels: 4
 DB: 30 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-671-687A-2 (1-3715)

Qy 1 MetSerSerGlyLeuTTPSerGlnGluValThrSerProTyrTTPGluArgIle 20
 Db 497 ATGAGTTCAGGCTTATGAGCCCAAGAAAAAGTCTTACCCCTACTGGGAAGCGGATT 556
 Qy - 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 Db 557 TTTTACTTGTCTTCTTCAAGATGCGAGGTACAGCAACAAACAAAGCTCTTAAA 616
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 Db 617 GTACCGAAGGGAAGTAGGACAGTATATTCAAGATCCTTCTGTGGGCGCATTCAGGATT 676

Sequence 290, App
 Sequence 290, App
 Sequence 613, App
 Sequence 613, App
 Sequence 8151, App
 Sequence 8151, App
 Sequence 2028, App
 Sequence 4185, App
 Sequence 6145, App
 Sequence 6145, App

QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 677 CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAAAATCTAGCAAACTCATGCA 736
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 737 GTTCTCTTTGTTGATGAANGGATGTTGTAGATTAATGAAGAATTCACAGATTACTT 796
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 797 TTGGCAATATACCAATTTGTGAGGAGAGGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT 856
QY 120 LysGlyLeuGluIleAspValGlyCysProValLysValClnLeuArgSerGlyGluGlu 139
DB 857 AAAGGCTCCAAATAGACTGGGCTCTCTGTGAAGATGACGTGAGATCTGGGGAAGAA 916
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 917 AATTTCTGGAGTTGTACGCTTCAGAGACCCTGTTAGCAGAGAGGACAGTCTCCGGA 976
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 977 ATATTCTTTGGAGTTGAAATGCTGGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 1036
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 1037 TACCAAGGAAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 1096
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 1097 AAGCTAGAACTCATAGAAGATGATGACCTGCAATGGAAGTGAATACGAGGTCCTGGG 1156
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 1157 GACACAATGCAGGTGCAACTTCTCTTTGGAAATAAATCTCCAGAGTTCTTTGAAGGGT 1216
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 1217 CGAGAAACAATAGAAATCTGGAACAGTTATATCTGTGTGTTTGGCCAGAAAGAAGC 1276
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1277 TTAGGATATTTGTTGGTGTGACATGGATACCCCTATTGGCAACTGGGATGGAAGATT 1336
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1337 GATGAGTGCANCTTTGTAGTTTGGCTGTGTTGAAAGTACAATTTCTATTGCACATCAAT 1396
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 1397 GATATCATCCAGAGAGTGTGACGCAAGAAAGGAGGCTCCCAAACTTGGCTTTATGTCA 1456
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1457 AGAGGTGTTGGGCAAAAGGTTTCATCCAGTCAATAAATAAACCAAGGCTACAGGATCACC 1516
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1517 TCAGACCCCTGGAAATAGAACAGATCTGAATATTTTATACCTTAAATGGGCTCTCTGTT 1576
QY 357 AspSerGlnProGlnSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1577 GACTCACACCAACATCCAAATCAAAAATACATGTTACATGATGAGTTGCAGAAAGAC 1636
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
DB 1637 CTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCCTTCCACCACCTCCAG 1696
QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1697 CCTCTCTGTGAACTCACTGACCACCGAGAACAGATTCCACTCTTTTACCATTTCAGTCTC 1756

QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1757 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGACCCAG 1816
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1817 TCTGTAATGGAAGAGCTAAACACTGCACCCGTCCAAGAGAGTCCACCCCTTGGCCATGCT 1876
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1877 CCTGGGAACCTCACATGCTGTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAGAACCTCTCT 1936
QY 477 PheTyrGlyValIleArgTTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1937 TTTCTATGGGTAATCCGTTGGATCGGTGAGCCACAGACTGAATGAAGTCTCGCTCGA 1996
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1997 CTGGAACCTGGAAGATGATGCTGCAGGCTGTACGATGGAACCTTCAGAGGACACTCGGTAT 2056
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 2057 TTCACCTGTGCCCTGAAGAAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCTGACTCT 2116
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 2117 AGTGTTCATCATTCAGCCGCTTCCCAATCAGATTGAGCGCTGTAATCTTTTAGCATTT 2176
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 2177 GGAAGCTACTTAAGTAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 2236
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 2237 TTGGAGATAATGATTGGGAAGAGAAAGGCATCCAGGCTCATTAACAATCTTTGTACTTA 2296
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2297 GACTCAACCTTATCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTTACTTAGA 2356
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2357 CCACAAAGAAAGAACCATGATGAGATATTTATAGTGAACCCCAAGAGCTACTGAGGACAGAA 2416
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 2417 ATTGTTAACTCTCGAATAATATGATATGTGTGTCACACAAAAATATGAACCTGAGG 2476
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 2477 AAAATACTTGAAGGTTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAAGATCCTCAG 2536
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2537 GAATTTCTTGAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTTGTGTAATAAGA 2596
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2597 TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG 2656
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2657 AAGTTGGGCTTCCCAATTCAGAGTGTGTAGAATGGTCTTTTATCAACAGATAACCTG 2716
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
DB 2717 AAATTTGCAGAGGACCATCATGCTCATATTATTCAGATGCTCGATTTTGGAAAAAGACTTT 2776
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAsp 776
DB 2777 AAACTATTTAAAAAATTTTCTCTCTCGAAATTAATAATAACAGATTTTACTTGAAGAC 2836
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796

```
Db 2837 ATCCGAGCAGTCCGCGATATGTGAGGGCTTGCATATGATGAGTGTAGAGATCTCTAC 2896
Qy 797 AspAspProAspIleSerAlaGlyIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2897 GACGATCCGGACATCTCAGCTGGAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2956
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2957 GTCCACCTTCATCCGGAAGAGGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAGAC 3016
Qy 837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 3017 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGAGATTATTGTCT 3076
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 3077 GTTCTCTGATAGAACACGACCTATGTCTTTGTGAGTATGGGAAGACGATTCT 3136
Qy 877 AlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 3137 GCCTGGCTCTCTTTGACAGCATGGCCGATCGGATGGTGGTCAAGATGGCTTCAACATT 3196
Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 3197 CCTCAAGTCACCCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 3256
Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 3257 CATTCCTTGACCTCCAGGAGATCCAAAGCTGTCACGAAGACTGCTTTGTGATGATAT 3316
Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3317 ATGTGATGTACCAGATGCCAACAATGAGTTGTACAAA 3355
```

RESULT 2

US-09-522-303-1293
; Sequence 1293, Application US/09522303
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE OF INVENTION: HUMAN MAMMARY EPITHELIUM LIBRARY
; TITLE REFERENCE: 1600.1086-001
; CURRENT APPLICATION NUMBER: US/09/522,303
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/123,393
; EARLIER FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 1353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 4661
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(4661)
; OTHER INFORMATION: n = A,T,C or G
US-09-522-303-1293

Alignment Scores:
Pred. No.: 0 Length: 4661
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 23 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-522-303-1293 (1-4661)

```
Qy 1 MetSerSerGlyLeuTyrPheSerGlnGluValThrSerProTyrTyrGluGluArgIle 20
Db 231 ATGAGTTTCAGGCTTATGGAGCCAAAGAAAAAGTCACCTTCAACCTACTGGGAAGAGCGGATT 290
```

```
Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db 291 TTTTACTTCTTCTTCAAGAATGACGGTGTACAGACAAACAAACACAAAAGCTCCTTTAAA 350
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 351 GTACCCAGAGGAAGTATAGACAGTATATTCAGATCGTTCCTGTGGGGCATTCAGAGATT 410
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 411 CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 470
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 471 GTTCTCTTGTGTTGATGAAAAAGGATGTTGTAGAGATAAATGAAAAATTTACAGAGATTACTT 530
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 531 TTGGCAATTTACCAATTTGTGAGGAGGGTTTCAGCTCTGTTAAAAACAGAAAACAGACTAAGT 590
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 591 AAAGGCTCCNAATAGACGTGGCTGTCTCTGCAAGTAGACAGCTGAGATCTGGGAAGAA 650
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 651 AAATTTCTGGAGTTGATGCTTTCAGAGGAGCCCTGTTTAGCAGAGAGACAGCTCTCCGA 710
Qy 160 IlePhePheGlyValGluLeuLeuGluGlyValArgGlyGlnGlyPheThrAspGlyVal 179
Db 711 ATATTTCTTGGAGTTGAAATTTGCTGGGAAGAGTCTGGTCAAGGTTTTCACCTGACGGGTG 770
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 771 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTGTCATTGGAC 830
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 831 AGCTAGAACTCATAGAGATGATGACACTGCACTTGGAAGTGATTACCGCAGGTCTCTGGG 890
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 891 GACACAAATGCAGGTGCAACTTCTCTTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT 950
Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 951 GGAGAAACAATAGAACTGGAACAGTTATATTCTGTGATGTTTTCAGGAGAAAAGAAAGC 1010
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPhe 278
Db 1011 TTAGGATATTTTGTGTTGGTGGACATGGATAACCTATTGGCAACTGGGATGGAAGATT 1070
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1071 GATGGAGTGCAGCTTTGTAGTTTTCGTGTGTTGAAAGTACAAATTTCTATTGCACATCAAT 1130
Qy 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
Db 1131 GATATCATCCAGAGAGTGTGACGCAAGAAAGAGGCTCCCAAACTTGCCTTTATTGTGCA 1190
Qy 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1191 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATATATAACCAAGGCTACAGGATCTACC 1250
Qy 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1251 TCAGACCTCGAAATAGAAAACAGACTCTGAATATTTTATACCTTAAATGGGCTCTTCGTT 1310
Qy 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp 376
Db 1311 GACTCACACCAATCCAAATCAAAAAATACATGGGTACATTGTATGAAGTTGCAGAGAC 1370
```


QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
DB 1371 CTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTTCCACCACCCTCCAG 1430
QY 397 ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1431 CCTCTCTCTGTGAACCTCACTGACCACCGAGACAGATTCCACTCTTTACCATTCAGTCTC 1490
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1491 ACCAAGATGCCCAATACCAATGAAGTATTTGGCCACAGCTCCACTTTCTCTGTGAGCCAG 1550
QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1551 TCTGTAATGGAAGAGCTAAACACTGACCCGCTCCAAAGAGAGCTCCACCTTGGCCATGCTT 1610
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1611 CCTGGGAACCTACATGGTCTAGAGTGGGCTCATTTGGCTGAAATTAAAGGAGAACCTCTCT 1670
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1671 TTCTATGGGTAAATCCGTTGGATCGGTACGCCACAGGACTGAATGAAGTGTCTCGCTGGA 1730
QY 497 LeuGluLeuLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1731 CTGGAACTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTAT 1790
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1791 TTCACCTGTGCCCTGAAAGAGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGTACTCT 1850
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1851 AGSTTTGCATCATTCAGCGGTTTCCATCAGATTGAGCGCTGTAACCTCTTTAGCATTT 1910
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 1911 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGGCG 1970
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1971 TTGGAAGATAATGATTGGGAAGAAAGAGCATCCAGGGTCAATACAAATTTCTTTACTTAA 2030
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2031 GACTCAACCTTATCTGCTTATTTGCTTTATGTTCTGTTCTGGACACTGTGTACTAGA 2090
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2091 CCCAAAGAAAGAACGATGTAGAAATATATAGTGAACCCCAAGAGCTACTGAGGACAGAA 2150
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 2151 ATTTGTTAATCTCTCAGAAATATAGTATGTTGTGCCACAAAATTTATGAACCTGAGG 2210
QY 657 LysIleLeuGluLysValGluAlaLaserGlyPheThrSerGluGluLysAspProGlu 676
DB 2211 AAAATCTTGAAGGTGGAGCTGATCAGGATTTTACCTCTGAAGAAAGATCTCTGAG 2270
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
DB 2271 GAATTTCTTGAATATCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAAAATAAGA 2330
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2331 TCAGCAGGTCAAAAGGTACAAAGATTGTACTTCTATCAAAATTTTATGGAAGAAATGAG 2390
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
DB 2391 AAAAGTTGGCGTTCCCAAAATTCAGCAGTTGTTAGAAATGTTCTTTTATCAACAGTAACCTG 2450
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756

DB 2451 AAATTTGCAGAGCACCATCATGTCGATTATTTCAGATGCTCGATTTCGAAAAGACTTT 2510
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2511 AAATCTTTAAAAAATTTTCTCTCTGGAATTAATAATATAACAGATTTACTTTGAAGAC 2570
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2571 ACTCCAGACAGTGGCGGATATGTGGAGGCTTGCATGTATGTAGTGTAGAGAAATGCTAC 2630
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2631 GAGATCCGAGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAACCTGCAACACTCA 2690
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2691 GTCCACCTTCATCCGAAGAGGCTGAATCATATAATAACCCAGTGTCACTTCCCAAGAC 2750
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 2751 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTTGCT 2810
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
DB 2811 GTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTCT 2870
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 2871 GCCTGCTCTTCTTTCAGACATGCGCGATCGGATGGTGTGTCAGATGGCTTCAACATT 2930
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
DB 2931 CCTCAAGTCAACCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 2990
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
DB 2991 CATTCCTTGAGCTCCAGGAGAAATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3050
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 3051 ATGTGCATGTACCAGAGTCCACATGAGTTGTACAAA 3089

RESULT 3
US-09-315-788-1729
; Sequence 1729, Application US/09315788 - Abandoned.
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: MLN98-19PM
; CURRENT APPLICATION NUMBER: US/09/315,788
; CURRENT FILING DATE: 1999-05-21
; EARLIER APPLICATION NUMBER: 60/086,455
; EARLIER FILING DATE: 1998-05-22
; EARLIER APPLICATION NUMBER: 60/132,067
; EARLIER FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1729
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4664)
; OTHER INFORMATION: n = A,T,C or G
US-09-315-788-1729

Alignment Scores: 0 Length: 4664
Pred. No.: 4983.00 Matches: 948
Score:

Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatches:	1
Query Match:	98.99%	Indels:	4
DB:	20	Gaps:	4
US-09-671-687A-3 (1-949) x US-09-315-788-1729 (1-4664)			
QY	1	MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle	20
Db	230	ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCCACCTTACCTGGGAAGAGCGGATT	289
QY	21	PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
Db	290	TTTTACTTCTCTTCAAGAAATGACGGTTACAGACAAACAAACAAAGCTCTCTTAA	349
QY	41	ValProLysGlySerIleGlyTyrIleGlnAspArgSerValGlyHisSerArgIle	60
Db	350	GTACCGAAGGGAATATAGACAGTATATTCAGATCGTCTGTGGGCAATTCAGGATT	409
QY	61	ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80
Db	410	CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAAAATTCAGAGCAACCTCATGCA	469
QY	81	ValLeuPheValAspGlu---AspValValGluLysAsnGlnLysPheThrGluLeuLeu	99
Db	470	GTTCCTCTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGATTACTT	529
QY	100	LeuAlaIleThrAsnCysGluGluAArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
Db	530	TTGGCANTTACCANTTTGTAGGAGAGGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT	589
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
Db	590	AAAGGCTCCAAATAGACGTGGCTGTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA	649
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
Db	650	AAATTTCTCGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGCTCCCGA	709
QY	160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal	179
Db	710	ATATTTCTTGGAGTTGAAATGCTGGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGT	769
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
Db	770	TACCAAGGNAACAGCTTTTTCAGTGTGATGAGATTGTGGCTGTTTGTGATTGGAC	829
QY	199	LysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	218
Db	830	AAGCTAGAACTCATAGAAATGATGACACTGCAATTCGAAAGTGATACGCGGCTCTGGG	889
QY	219	AspThrMetGlnValGluLeuProProLeuGluLysAsnSerArgValSerLeuLysGly	238
Db	890	GACAAATGAGGTGCAACTTCCTCTTGGAAATTAACCTCCAGAGTTTCTTTGAAAGGTT	949
QY	239	GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer	258
Db	950	GGAGAAACAATAGAACTGGAACAGTTATATCTGTGATGTTTGGCAGGAAAGAAAGC	1009
QY	259	LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe	278
Db	1010	TTAGGATATTTTGTGTGGTGGACATGATTAACCTATTGGCAACTGGGATGGAAGATT	1069
QY	279	AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn	297
Db	1070	GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTTGAAAGTACAATTTATTTGCAATCAAT	1129
QY	298	AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer	317
Db	1130	GATATCATCCAGAGATGTGACGAGAAAGAGGCTCCCAAACTTGGCTTTATGTCA	1189
QY	318	ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr	337

Db	1190	AGAGGTGTTGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC	1249
QY	338	SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal	356
Db	1250	TCAGACCTTGGAAATAGAAACAGATCTGAATTTATTTATACCTTAAATGGTCTTCTGTT	1309
QY	357	AspSerGlnProGlnSerLysAsnThrTyrIleAspGluValAlaGluAsp	376
Db	1310	GACTCAACACCAATCAAAATCAAAATAATACATGTTGATGAAAGTTGCAAGAAC	1369
QY	377	ProAlaLysSerLeuThrGluLysSerThrAspPheAspArgSerSerProLeuGln	396
Db	1370	CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCCGTTCTTCCACCACCTCCAG	1429
QY	397	ProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu	416
Db	1430	CCTCTCTCTGTGAACCTCACTGACCACCGAGAACAGATTCACCTTTTACCATTTCAGTCTC	1489
QY	417	ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln	436
Db	1490	ACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG	1549
QY	437	SerValMetGluLysLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro	456
Db	1550	TCGTAAATGGAAGAGCTAAACACTGCACCCGCTCCAGAGAGTCCACCTTGGCCCATGCT	1609
QY	457	ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro	476
Db	1610	CCTGGGAATCCTCATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCT	1669
QY	477	PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly	496
Db	1670	TTCTATGGGGTAATCCGTTGGATCGGTGAGCTGAGCCACAGGACTGAATGAAGTCTCGCTGA	1729
QY	497	LeuGluLeuGluLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516
Db	1730	CTGGAATGGAAGATGAGTGTGAGGCTGTACGAGTGAACCTTTCAGAGGACTCGGTAT	1789
QY	517	PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer	536
Db	1790	TTCACTGTGCTTGAAGAGGCGCTGTTGTGAACTGAAGAGCTGACGCTGACTCT	1849
QY	537	ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe	556
Db	1850	AGGTTTGCATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTGTAATCTTTAGCATTT	1909
QY	557	GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly	576
Db	1910	GGAGGCTACTTAAAGTAGAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGGC	1969
QY	577	LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu	596
Db	1970	TTGGAGATTAATGTTGGAAAGAAAGAGGATCCAGGCTCATTTACAATTTCTTGTACTTA	2029
QY	597	AspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArg	616
Db	2030	GACTCAACCTTATCTGCTTATTTGCTTTTGTGTTCTGTTCTGAGACACTGTTACTTGA	2089
QY	617	ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
Db	2090	CCCAAGAAAGAAAGACGATGTAGAAATATATATAGTGAACCCCAAGAGCTACTGAGACAGAA	2149
QY	637	IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg	656
Db	2150	ATTGTTTAATCTCTGAGATATATGATATGTTGTGTCACAAATTTATGAACTGAGG	2209
QY	657	LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
Db	2210	AAATACTTGAAGGAGGCTGATCAGGATTTTACCTCTGAAAGAAAGAAAGATCTCTGAG	2269
QY	677	GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg	696
Db	2270	GAATTTCTGAAATATCTGTTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAGA	2329

QY 697 SerAlaGlyGlnIysValGlnAspCysTyrPheTyrGlnIlePheMetGluIysAsnGlu 716
Db 2330 TCAGCAGGTCAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGGAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCGTCCCAACATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCACCATCATGCTGTGATTTATTCAGATGCTCGATTTGMAAAGACTTT 2509
QY 757 LysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuAsp 776
Db 2510 AAACATATTTAAAAAATTTTCTCTCTGGAATTAATAACAGATTTACTTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCCGATATGTGGAGGCTTGCAATGTATGATGTAGAGAAATGCTAC 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTGTAAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGGCTGAAATCATAAATATAAATCAACAGTGTCTTCCCAAAGAC 2749
QY 837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGAGCTGGAGACACGCTGCATCCCTTGCAGAAATATGGAGTTATTGTCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
Db 2810 GTTCTCTGATAGAAACAGACCACTATGTGTTTTGTGAGTATGGAGGACGATTTCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGGCTCTCTTTGACAGCATGGCGGATCGGATGGTGGTTCAGAAATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCACCCCATGCCAGAGTAGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGGACTCCAGGAGATCCCAAGGCTGTGCACGAGACTGCTTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACAGAGTCCAAACATGAGTTGTACAAA 3088

RESULT 4

US-09-315-788A-1729
; Sequence 1729, Application US/09315788A — *Abdn*
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: 1600, 1019-002
; CURRENT APPLICATION NUMBER: US/09/315, 788A
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 09/315, 788
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/086, 455
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/132, 067
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1729

; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(4664)
; OTHER INFORMATION: n = A, T, C or G
US-09-315-788A-1729

Alignment Scores:

Pred. No.: 0 Length: 4664
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 20 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-315-788A-1729 (1-4664)

QY 1 MetSerSerGlyLeuTyrPheSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACCTTCACTTCTGGAAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCAGGTTACAGACAAACAAACACAAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCCAGAGGAGATATAGACAGTATATTCAGATCGTCTCTGGGGGCATTCAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLysLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTGCAAAAGCGCAAGAAAAATCAGATTGGATTAATAATCTTAGACCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTGTGATGAAAAAGGATGTTGTAGAGATAAATGAAAAAGTTTCACAGAGTTACTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTACCAATTTGTGGAGAGAGGTTACAGCTGTTTAAATAACAGAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTCGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATCTTTGGAGTTGAATTTGCTGGAAGAGGTCGTGCTCAAGGTTTCACAGCGGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829
QY 199 LysLeuLeuLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCATTTGAAAGTATTACCGAGGCTCTGGG 889
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACAAATGCAGGTCGAACCTCTCTCTTGGAAATAAACTCCAGAGTTCTTTTGAAGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 950 GGAGAAACATAGAAATCTGGAACAGTTATATTCGTGATGTTTTCGCCAGAAAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278

Db 1010 TTAGGATATTTTGGTGGACATGGATACCCTATTGGCAACTGGGATGGAAGATTT 1069
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1070 GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTTGAAAGTACAATTCTATTGGACATCAAT 1129
QY 298 AsnIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
Db 1130 GATATCATCCAGAGAGTGTGACGAGGAAGAGGAGGCTCCCAAACTTGCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1250 TCAGACCTCGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
Db 1310 GACTCAACACCAATCCAAATCAAAATAATACATGATGATGATGATGATGATGATGATGAT 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
Db 1370 CCTGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTTCTTCACCCACCTCCAG 1429
QY 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CCTCTCTCTGTGNACTCATGACCACCGAGACAGATTCACCTCTTTACATTCAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCAATACCAATGAAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG 1549
QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1550 TCTGTAATGGAGAGCTAAACACTGCACCGCTCCAAAGAGAGTCCACCTTTGGCCATGCCT 1609
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CCTGGGAACCTCATGGCTTAGAAGTGGGCTCAATGGCTGAAAGTAAAGGAAACCTCTCT 1669
QY 477 PheTyrGlyValIleArgThrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCTATGGGGTAATCCGTTGGATCGGTGACCGACAGGACTGAATGAAGTCTCGTGGTA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACTGGGAAGTGAAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTAT 1789
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCTGAAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCGCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGGTTTGCATCATGTGACCGGTTTCCAATCAGATTGAGGCGCTGAACCTTTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATTGGGAAGAGAAAGGCAATCCAGGCTCATTTACAAATCTCTGTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTGCTTTTGTCTTTTAGTTCTGTCGACACTGTGTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAGAGAAAGAACCATGTAGTAATATTATAGTGAACCCCAAGAGCTACTGAGACAGAA 2149

QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAATCTCTGAGAATATATGATGATGTGTGCCACAAAAATATGAAACTGAGG 2209
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTTGAAAGGTGGAGGCTGCATCAGGATTTTACCTCTCGAAGAAAAAGATCCCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTCAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTTAAAAATAAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAGATTGTACTTCTATCAAAATTTTATGGAAGAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCGTTTCCCACAATTCAGCAGTTGTTTAAAGATGGTCTTTTATCAACAGCTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCCACCATCATGTCGATTAATTCAGATGCTCGATTTGGAAAAAGACTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATTTTAAAAAAATTTTCTCTCTCGAAATTAATAATAACAGATTTTACTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCCGGATATGTGGAGGCTTGCATGATGATGATGATGATGATGATGAT 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCCGACACTCAGCTGGAATAATCAAGCAGTTTGTAAAAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACTTCATCCGAAGAGGCTGAATCATATATATACCCAGTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACACGCTGCATCCCTTCCAGAAATATGGAGTTATTGCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCTGCTCTCTTTTTCAGCATGCGCGATCGGAGTGGTGTGAGATGTCTCTGGAAGACTT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGGACTCCAGGAAATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACCAGAGTCCAAATGAGTTGTGACAAA 3088

RESULT 5

US-09-783-514-1729
; Sequence 1729, Application US/09783514
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; HUMAN ASTROCYTE LIBRARY

- Asn

FILE REFERENCE: 1600.1019-002
 CURRENT APPLICATION NUMBER: US/09/783,514
 CURRENT FILING DATE: 2001-02-14
 PRIOR APPLICATION NUMBER: 09/315,788
 PRIOR FILING DATE: 1999-05-21
 PRIOR APPLICATION NUMBER: 60/086,455
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/132,067
 PRIOR FILING DATE: 1999-04-30
 NUMBER OF SEQ ID NOS: 2346
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1729
 LENGTH: 4664
 TYPE: DNA
 ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(4664)
 OTHER INFORMATION: n = A,T,C or G
 US-09-783-514-1729

Alignment Scores:
 Pred. No.: 0 Length: 4664
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 34 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-783-514-1729 (1-4664)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
 DB 230 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCACTCCCTACTGGGAAGCGGATT 289
 QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
 DB 290 TTTTACTTCTCTTCCAGAATGCAGCGTTACAGACAAAACAAACACAAAGCTCCTTAA 349
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
 DB 350 GTACCGAAGGGAAGTAGACAGTATATTCAGATCGTCTGTGGGCATTCAGGATT 409
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
 DB 410 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATTCTAGAGCAACCTCATGCA 469
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
 DB 470 GTTCTCTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGATTACTT 529
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 530 TTGGCAATTTACCAATTTGAGGAGAGGTTGAGGCTGTGTTTAAACACAGAACACACTAAGT 589
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 590 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 649
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 650 AAATTTCTCGAGTTGATCGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGCTCCCGA 709
 QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
 DB 710 ATATTCTTTGGAGTTGAATTTGCTGGAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 770 TACCAGGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTTCATTGGAC 829
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218

DB 830 AAGCTAGAACTCATAGAAGATGATGACACTGCATTGGAAAGTGATTACGAGGTCCTGGG 889
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 890 GACACAATGCAGGTCGAACTTCTCTTGGAAATAAACTCCAGAGTTCTTTGAAGGTT 949
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 950 GGAGAAACAATAGAACTCTGGAAACAGTTATATTCTGTGATGTTTTCGCCAGGAAAAAGAAC 1009
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1010 TTAGGATATTTTGTGTGGTGGACATGATAACCTATTGGCAACTGGATGGAGATT 1069
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1070 GATGGAGTCAGCTTTTGTAGTTTTCGCTGTGTGAAAGTACAAATCTATTGCACATCAAT 1129
 QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
 DB 1130 GATATCATCCAGAGAGTGGACGAGAAAGGAGGCTCCCAACTTTGCCTTTATGTCA 1189
 QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 1190 AGAGGTGTTGGGACAAAGGTTCTATCCAGTCATATAAACCACAAAGGCTACAGGATCTACC 1249
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 DB 1250 TCAGACCTCGAAATAGAAACAGATCTGAATTAATTTATACCTTAAATGGGTCTCTGTT 1309
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
 DB 1310 GACTCAACACCAATCCAAATCAAAATAATACATGATGATGATGATGATGATGATGATGAT 1369
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGln 396
 DB 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTACCCGCTTCTTACCACTCCAG 1429
 QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 DB 1430 CCTCTCTGTGAACCTCACTGACCCAGACAGACAGATTCACACTCTTTACCATTCAGTCTC 1489
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 1490 ACCAAGATGCCAATCAATGGAAGTATTTGGCCACAGTCCACTTCTCTCTCAGCCAG 1549
 QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 1550 TCTGTAATGGAAGAGCTTAAACACTGACCCGTCACAGAGAGTCCACCCCTTGGCCATGCT 1609
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 1610 CTGTGGNACTCATGCTTAGAAGTGGGCTCATTTGCTGAGTTAAGGAGAACCTCTCT 1669
 QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 1670 TTCTATGGGTAAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTCTCGCTGGA 1729
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 DB 1730 CTGGAACCTGAAGATGATGTGAGGCTGTACGGATGGAACCTTCAGAGGACCTCGGTAT 1789
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspSer 536
 DB 1790 TTTACCTGTCCCTGGAAGAGGCGCTGTTTGTGAAACTGAAAGAGCTGAGGCGCTGACTCT 1849
 QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
 DB 1850 AGTTTGCATCATTTGACGCGGTTTCCAAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
 QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGly 576
 DB 1910 GGAGGCTTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGCAAAAGAGAGGC 1969

```
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATTGGGAAGAAGAGCATCCAGGGTCATTACAAATCTTGTACTTA 2029

QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTATTGCTTTAGTTCTGTTCTGGACACTGTGTTACTTAGA 2089

QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAAGAACCATGTAGATAATTATAGTGAACCCCAAGACTACTGAGACAGAA 2149

QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAATCCTCGAATATATGATATGTTGTGCCACAAAAATATGAAACTGAGG 2209

QY 657 LysIleLeuGluLysValGluAlaLaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCCTGAG 2269

QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTGNATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTAAAAATAAGA 2329

QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGGAAGAAAATGAG 2389

QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleLeuSerAsnLeu 736
Db 2390 AAAGTTGGGGTTCGCCAATTCACAGATGTTGTTAGAATGGTCTTTTATCAACAGTAACCTG 2449

QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCACATCATGCTCTGATTTATTCAGATGCTCGATTTGGAAGAACTTT 2509

QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGAC 2569

QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTGCAGGATATGTGGAGGCTTGCATATGATGATGATGATGATGATGATG 2629

QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGGACATCTCAGCTGGAAATAATCAAGCAGTTTGTAAAAACCTGCAACACTCAA 2689

QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGCTGATCATATATATATACCATGCTGCTCTCCCAAGAC 2749

QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGACACAGCGCTGCATCCCTTGCAGATAATGAGATTTATTGCT 2809

QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAAACCAACCACTATGTTGCTTTTGTGAAGTATGGAAGAGCAGATTCT 2869

QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGGCTCTTCTTTGACATGCGCGATCGGATGGTGGTTCAGAAATGGCTTCAACATT 2929

QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCATGCCAGAGTAGGAGTACTTTGAAGATGCTCTCTGGAAGACCTG 2989

QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGGACTCCAGAGATCCCAAGGCTGTGCAGAGACTGCTTGTGTGTCATAT 3049
```

```
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACCAGAGTCCAAATGAGTTTGTACAAA 3088

RESULT 6
US-09-783-514A-1729 - Adm
; Sequence 1729, Application US/09783514A
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: 1600.1019-002
; CURRENT APPLICATION NUMBER: US/09/783,514A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/315,788
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/086,455
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/132,067
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1729
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(4664)
; OTHER INFORMATION: n = A,T,C or G
US-09-783-514A-1729

Alignment Scores:
Pred. No.: 0 Length: 4664
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 34 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-783-514A-1729 (1-4664)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTTCAGCTTATGGAGCCCAAGAAAAGTCACTTCACTTCTGGAAGACGGGATT 289

QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACAGACAAACAAACAAAGCTCCTTAAA 349

QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCAAGGGAAGTAGTAGACAGTATATTTCAAGATCGTTCTGTGGGGCATTTCAAGGATT 409

QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTGCAAGCAAGCAAGAAAATCAGATTGGATTAAAAATTTCTAGACCAACCTCATGCA 469

QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTTCTGATGAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGATTACTT 529

QY 100 LeuAlaIleThrAsnCysGluLysArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTAACCAATTTGTGAGGAGAGGTTTCCAGCTGTGTTAAAAACAGAAAACAGTAACT 589

QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTTCCAAATAGACGTTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGAAGAA 649

QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
```

Db 650 AAATTTCTCGAGTTGTACGCTTCAGAGGACCCCTGTGTAGCAGAGGACAGTCTCCGGA 709
Qy 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTCTTTGGAGTTGAATTCGTGGAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGGCGTGTGTGTTCATTGGAC 829
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCAATTGGAAAGTGATTACCGAGCTCTGGG 889
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACAATGCAGTTCGAATCTCTCTTTTGGAAATAAATCCAGAGTTTCTTTTGAAGTT 949
Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 950 GGAGAAACAATAAGATCTGGAACAGTTATTTCTGTGATGTTTTCAGGAAAGAAAGC 1009
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 1010 TTAGGATATTTTGGTGTGACATGGATACCCCTATTGGCAACTGGGATGGAAGATT 1069
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1070 GATGAGTGCAGCTTTGTAGTTTGGTGTGTGAAAGTACAATTTCTATTGCACATCAAT 1129
Qy 298 AsnIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
Db 1130 GATATCATCTCCAGAGAGTGTGACGAGAAAGAGGCGCTCCCAAACTTGCTTTATGTCA 1189
Qy 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1190 AAGGTGTGTGGGACAAAGGTTTCATCCAGTCATATAAACCAAGGCTACAGATCTACC 1249
Qy 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1250 TCAGACCTCGGAAATAGAAACAGATCTGAATTTATTTTATACCTTAAATGGGTTCTGTT 1309
Qy 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
Db 1310 GACTCACACCAACCAATCCAAATCAAAAAATACATGTTACATTTGATCAAGTTGCAGAAGAC 1369
Qy 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
Db 1370 CCTGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTTCTTCACCAACCACTCCAG 1429
Qy 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CCTCTCTGTGAACTCACTGACCCAGGAGACAGATTCCACTCTTTACATTCACTCTC 1489
Qy 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLysSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG 1549
Qy 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1550 TCTGTAATGGAGAGCTAAACACTGCACCCGCTCCAAAGAGAGTCCACCTTGGCCATGSCCT 1609
Qy 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CTGGGAATCTACATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT 1669
Qy 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCTATGGGTAAATCCGTTGGATCGGTGACGCCACAGGACTGAATGAAGTGTCTCGTGA 1729
Qy 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACGTGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTTCAGAGGCACTCGGTAT 1789

Qy 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCCCTGAGAAAGGCGCTGTTTGTGAAACTGAGAGCTGCAGCCCTGACTCT 1849
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGTGTTCATCATTCAGCCGGTTCACATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
Qy 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 1969
Qy 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATTGGGAAGAAAGGATCCAGGCTCATTTACAAATTCCTTTGTACTTA 2029
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTGA 2089
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAGAACCGATGTAGAATATTATAGTGAACCCCAAGAGCTACTGAGCAGAA 2149
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAAATCTCTGAGATATATGATATGTTGTGTGCCACAAAAATTTATGAAACTGAGG 2209
Qy 657 LysIleLeuGluLysValGluAlaAspGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATCTTTGAAAAGTGGAGGCTGCATCAGGATTTTACCTCTGAAAGAAAAAGATCCTGAG 2269
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATTTCTTGATATTTCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAAAATAAGA 2329
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGGAAAAAATGAG 2389
Qy 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCGTTCCTCCCAATTCAGCAGTTGTTAGAATGGTCTTTTATCAACAGTAACCTG 2449
Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGGACCATCATGTTCTGATTATTTCAGATGCTCGATTTGGAAAAAGACTTT 2509
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAATTTTAAAAAATTTTCTCTCTGGAATTAATAATATAACAGATTTACTTTGAAGAC 2569
Qy 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTGCAGGATATGTGGAGGCTTGCATATGATGATGAGTGTAGAGAATGCTAC 2629
Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GAGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAAAATCTGCAACACTCAA 2689
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACTTCATCCGAAGAGGCTGAATCATATAATATAACCCAGTGTCTACTTCCCAAGAC 2749
Qy 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACACGCTGCATCCCTGCGAGAAATATGGAGTTATTGCT 2809
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTCT 2869

```

QY 877 AlaTrrLeuPheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGGCTCTTTTACAGCATGCGCGATCGGATGGGTGAGAAATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCACCCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTGGACTCCAGGAGATCCAAAGCGTGTGCACGAAGACTGCTTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACAGAGTCCAAATAGTGTGTACAAA 3088

RESULT 7
US-09-396-087-4309 - Abdn
; Sequence 4309, Application US/09396087
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: PULMONARY ARTERY SMOOTH MUSCLE LIBRARY
; FILE REFERENCE: MLN98-39pM
; CURRENT APPLICATION NUMBER: US/09/396,087
; EARLIER FILING DATE: 1999-09-14
; EARLIER APPLICATION NUMBER: 60/100,260
; EARLIER FILING DATE: 1998-09-14
; EARLIER APPLICATION NUMBER: 60/107,226
; EARLIER FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: 60/131,810
; EARLIER FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5220
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4309
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-396-087-4309

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 21 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-396-087-4309 (1-4668)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
Db 230 ATGAGTTCAAGCTTATGGAGCCCAAGAAAGATCATCTTCCCTTCTGGGAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGCTTCTCAAGAAATGACGGTTTACAGACAAACAAACAAAGCTCTCTTAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisArgIle 60
Db 350 GTACCGAAGGAGATGATAGACAGATATATTCAAGATCGTTCTGTGGGGCATTCAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
Db 410 CCTTCTGCAAAAGCAAGAAATATAGATTGGATTAAATTTCTAGAGCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeu 99

```

```

Db 470 GTTCTCTTTGTTGATGATAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 529
QY 100 LeuAlaIleThrAsnCysGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTTACCAATTTGTGAGAGAGGTTTTCAGCTCTGTTTAAACAGAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGCGCTCCAAATAGACGTGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTCTGGAGTTGTACGCTTCAGAGACCCCTGTAGCAGAGAGGACAGTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuGluGlyValArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTTCTTTGGAGTTGAATTTGCTGGAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTTGTGGCTGTTTGTTCATTGGAC 829
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCAATGGAAAGTGAATACGCGAGGTCCTGGG 889
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACATGCGAGTTCGAATCTCTCTTTGGAAATAAATCCAGAGTTTCTTTGAAAGGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 950 GGAGAAACAAATAGAATCTGGAACAGTTATATCTGTGATGTTTTCAGGAGGAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 1010 TTAGGATATTTTGTGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuIleAsn 297
Db 1070 GATGGAGTGCAGCTTTGTGATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
Db 1130 GATATCATCCAGAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1190 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCAATAATAAACCAAGGCTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1250 TCAGACCTCGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
Db 1310 GACTCACACCAATCCAAATCAAAAAATACATGATGATGATGATGATGATGATGATGATGATGAT 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
Db 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCACCACTCCAG 1429
QY 397 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CCTCTCTCTGTGAACCTACCTGACCCAGAGAGATTCACACTCTTTTACCATTTCAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCAATACCATTGGAGATTTGGCCACAGTCCACTTCTCTGTACGCCAG 1549
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1550 TCTGTAATGGAAGAGCTAAACACTGCACCGCTCAAGAGAGTCCACCGCTTGGCCATGCCT 1609

```



```
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CCTGGAACTCAGATGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCCT 1669
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCTATGGGGTAACTCGTGGATCGGTACGCCACAGGACTGAATGAAGTGTCTCGCTGGA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACTCGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTAT 1789
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCCCTGAAGAAGGCGCTGTTGTGAAACTGAAGAGACTGCAGGCGCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGTTTGGCATCATTTGCAGCGGTTTCCAATCAGATTGAGGCGCTGAACCTCTTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATTAATGATTCGGAGAGAAAGGCAATCAGGGCTCATTAACAATCTCTGTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTAATCTGCTATTGCTTTAGTTCTGTTCTGGACACTGTGTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAGAACGATGTAGAAATATTATAGTGAACCCAGAGACTACTGAGACAGAA 2149
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAATCCTGAGATATATGATGATATGCTGTGCCACAAAAATTTATGAACCTGAG 2209
QY 657 LysIleLeuGluLysValGluAlaIleSerGlyPheThrSerGluLysAspProGlu 676
Db 2210 AAAATACTTTGAAAAGTGGAGGCTGCATCAGGATTTACCTCTCGAAGAAAAGATCTCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTGAATATTCTGTTTCATCATATTTTAAAGGGTGAACCTTTGCTAAAAAATAAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAGATTGTACTTCTATCAATTTTATGGAAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAAGTTGGGTTCCCAACAATTCAGCAGTTGTTAGAAATGGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAAATTTGCAGAGGCACCATCATGCTGATTATTTCAGATGCTCGATTTTGGAAAAAGACTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATATTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTACTTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCGGATATGTGGAGGGCTTGCAATGATGATGATGATGATGATGATGAT 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2689
```

```
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGGCTGAATATAAAATATAACCCAGTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGGCAGAAATATGGAGTTATTGCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAAACAGCCACTATGTGTTGCTTTGTGAAGTATGGAGAGACGATTCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCTGGCTCTCTTTTGCACATGGCGGATGGTGGTCAAGATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCATGCCAGAAAGTAGGAGAGTACTTTGAAGATGTCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGGACTCCAGGAGAAATCCAAGGCTGTGCACGAGACTGCTTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACCAGAGTCCAAACAATGAGTTGTACAAA 3088

RESULT 8
US-09-396-970-7559
; Sequence 7559, Application US/09396970 -> Node
; GENERAL INFORMATION:
; APPLICANT: Kingsbury, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: MIXED LYMPHOCYTE LIBRARY
; FILE REFERENCE: MLN98-40Pa
; CURRENT APPLICATION NUMBER: US/09/396,970
; CURRENT FILING DATE: 1999-09-14
; EARLIER APPLICATION NUMBER: 60/100,293
; EARLIER FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 8756
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7559
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-396-970-7559

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 0 948.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 21 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-396-970-7559 (1-4668)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
Db 230 ATGAGTTTCAGGCTTATGGAGCCAAAGAAAAAGTCACTTACCCCTATGCGAAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGATGCGAGGTTACAGACAAACAAACAAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
```


Db 350 GTACCGAAGGAGTATAGACAGTATATTCAAGATCGTTCTGTGGGCAATTCAGGATT 409
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAATTCAGAGCAACCTCATGCA 469
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGATTACTT 529
Qy 100 LeuAlaIleThrAsnCysGluGluAtrgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTTACCAATTTGTGAGAGAGGTTTCAGGCTGTTTAAACAGAAACAGACTAAGT 589
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTCGAGTTGTAGCGTTTCAGAGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 709
Qy 160 IlePhePheGlyValGluLeuLeuGluGluGlyAtrgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATCTTTGGAGTTGAATTTCTGGAGAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGGCGTGTGTTGTGATTTGGAC 829
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAAGATGATGACCTGCATTTGGAAAGTGATTACGCAGGTCCTGGG 889
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACAATGCAGTCGAATCTCTCTCTTGGAAATAAATCCAGAGTTCTTTTGAAGGTT 949
Qy 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
Db 950 GGAGAAACATAGAACTCGAACAGTTATATCTGTGATGTTTTCAGGAAAGAAAGC 1009
Qy 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
Db 1010 TTAGGATATTTGTTGGTGTGGACATGATGATAACCTATTGGCACTGGGATGGAAGATT 1069
Qy 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
Db 1070 GATGGAGTCAGCTTTGTAGTTTTCGTGTGTTGAAAGTACAATTTCTATTGCAATCAAT 1129
Qy 298 AspIleIleProGluSerValThrGlnGluAtrgArgProProLysLeuAlaPheMetSer 317
Db 1130 GATATCATCCAGAGAGTGTGACGCAAGAAAGAGGCCCTCCCAAACTTGCCTTTATGTCA 1189
Qy 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
Db 1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACC 1249
Qy 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
Db 1250 TCAGACCCCTGGAATAAGAAACAGATCTGAATATATTTATACCTTAAATGGGCTCTCTGTT 1309
Qy 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
Db 1310 GACTCAACACCACCAATCCAAATCAAAAAATACATGTTACATTTGATGAAGTTGCAGAGAC 1369
Qy 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
Db 1370 CCTGCAAAATCTTACAGAGATATCTACAGACTTTGACCGTTCTTCACCAACCACTCCAG 1429
Qy 397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CCTCCTCTGTGAACCTCACTGACCAACCGAGAACAGATTCCACTCTTTTACCATTTCAGTCTC 1489

Qy 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG 1549
Qy 437 SerValMetGluGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1550 TCTGTAAATGGAAGAGCTAAACACTGCACCCGCTCCAAGAGAGTCCACCTTTGGCCATGCT 1609
Qy 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CTTGGGAATCTCAGATGGTCTAGAACTGGGCTCAATTTGGCTGAAGTTAAGAGAACCTCTCT 1669
Qy 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTTCTATGGGGTAAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTCTCGCTGGA 1729
Qy 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACTGGAAGATGAGTGTGAGGCTGTGACGATGGAACTTTCAGAGGCACTCGGTAT 1789
Qy 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
Db 1790 TTTCACTGTGCTCTGAAGAGGCGCTGTTTGTGAACTGAAGAGCTGACGCGCTGACTCT 1849
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGSTTTGTCATCATTTGCAGCCGGTTTCCATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
Qy 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAAATCTCCACCAAAAATGGAAAAAGAGGC 1969
Qy 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATTCGGAGAGAAAGGCAATCCAGGCTCATTTACAAATTTCTGTACTTA 2029
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATTTCTGCTTATTTGGCTTTAGTTCTGTCTGGACACTGTGTACTTAGA 2089
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAAGAACGATGTAGAATATATAGTGAACCAACCAAGAGCTACTGAGACAGAA 2149
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTTGTTAAATCCTCTGAGAATATATGATATGTGTGTGCCACAAAATTTATGAAACTGAGG 2209
Qy 657 LysIleLeuGluLysValGluAlaLaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAATTAATTTGAAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCTCTGAG 2269
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATTTCTTGAATATTTCTGTTTCATCATATTTTAAAGGTGAGAACCTTTGCTAAAAATAAGA 2329
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAAGATTGTACTTCTATCAAAATTTTATGGAAGAAAATGAG 2389
Qy 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGGCTTCCCAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACTGT 2449
Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGGCAACCATCTGCTGATTTATTTATTCAGATGCTCGATTTGGAAGACTTT 2509
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAATTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGAC 2569

QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2570 ACTCCAGAGAGTCCCGGATATGTGGAGGCTTGCAATGATGAGTGTAGAGTAATGCTAC 2529
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2630 GACGATCCGGACATCTCAGCTGGAAATAATCAACGAGTTTGTGTAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2690 GTCCACCTTTCATCCGAAGAGGCTGAATCATATAATAAACCAGTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 2750 TTACCCGACTGGGACTGGGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTGGCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
DB 2810 GTTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGGAAGGAGGATTCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 2870 GGCTGCTCTTCTTTCACAGCATGGCGCATCGGATGGTGGTCAGAAATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
DB 2930 CCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
DB 2990 CATTCCTTGAGCTCCAGGAGATCCAAAGCTGTGCACGAGACTGCTTTGTGATGATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 3050 ATGTGATGTACCAGAGTCCAAACATGAGTTGTACAAA 3088

RESULT 9

US-09-397-424-5019
; Sequence 5019, Application US/09397424 — *Abdn*
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: MLN98-45pm
; CURRENT APPLICATION NUMBER: US/09/397,424
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,469
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/106,454
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,252
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/132,100
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5379
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5019
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-397-424-5019

Alignment Scores:

Pred. No.:	0	Length:	4668
Score:	4983.00	Matches:	948
Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatches:	1
Query Match:	98.99%	Indels:	4

DB: 21 Gaps: 4
US-09-671-687A-3 (1-949) x US-09-397-424-5019 (1-4668)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 230 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCACTCCCTACTGGGAAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
DB 290 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACAGACAAACAAACACAAAAGCTCCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 350 GTACCAAGAGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCGCATTCAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
DB 410 CCTTCTGCAAAAGCAAGAAAATCAGATTGGATTAAAATTTCTAGAGCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
DB 470 GTTCTCTTTCATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 530 TTGGCAATTTACCAATTTGTAGAGAGAGGTTTCAGGCTGTTTAAAAACAGAAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 590 AAAGGCTTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 650 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGAGACAGTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179
DB 710 ATATTTCTTGGAGTTGAAATTTGCTGGAAGAGGTCGTGCTCAAGGTTTTCACCTGACGGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 770 TACCAAGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTTCATTTGAC 829
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 830 AAGCTAGACTCATAGAGATGATGACATGCACTTGAAGAGTATTACGAGGTCCTGGG 889
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 890 GACACAAATGCAGGTCGAACTTCTCTTGGAAATAAATCCAGAGTTTCTTTGAAGGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 950 GGAGAAACAAATAGAAATCTGGAACAGTTATATCTGTGATGTTTTCGCCAGGAAAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPhe 278
DB 1010 TTAGGATATTTTGTGTTGGACATGGATAACCTATTGGCAACTGGATGGAGATT 1069
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1070 GATGGAGTCAGCTTTGTAGTTTTCGCTGTTGAAAGTACAATTTCTATTTCACATCAAT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 1130 GATATCATCCAGAGAGTGTGACGACAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATATAAACCAGGCTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356

Db 1250 TCAGACCCCTGGAATAGAACAGATCTGAATATTTTATACCTTAATGGGTCTTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrTrpIleAspGluValAlaGluAsp 376
Db 1310 GACTCAACAACCAATCCAAATCAAAAAATACATGATGATGAAAGTGCAGAGAC 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
Db 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCACCACTCCAG 1429
QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
Db 1430 CCTCTCTCTGTGAACCTACAGACCGAGACAGATTCACCTCTTTACCATTCAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
Db 1490 ACCAAGATGCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTTCTCTGTGAGCCAG 1549
QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
Db 1550 TCTGTAAATGGAAGAGCTTAAACCTGACCCCGCTCAAGAGAGTCCACCCCTTGGCCATGCCT 1609
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
Db 1610 CCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTAAAGGAGAACCCCTCT 1669
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
Db 1670 TTCATATGGGGTAAATCCGTGTGGATCGGTGAGCCACAGGACTGAATGAGTGTCTCGTGA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
Db 1730 CTGGAACCTGGAAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGGCACTCGGTAT 1789
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspSer 536
Db 1790 TTCACCTGTGCTGGAAGAGGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCGCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGGTGTTCATCATTTGACGCGGTTTCCATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATAATGATTGGGAAGAGAAAGGCATCCAGGGTCAATACAAATCTCTGTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTTGCTTTAGTTTGTCTGTTGACACTGTGTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAGAAAGAACGATGATAGATATATATAGTGAACCCAGAGCTACTGAGACAGAA 2149
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAATCCTCTGAGATATATGATATGTGTGTGCCACAAAAATATGAACTGAGG 2209
QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTGAAGAGGTGGAGGCTGCATCAGATTTTACCTCTGAAGAAAAAGATCCCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTGAATATCTCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTTAAAAATAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716

Db 2330 TCAGCAGGTCAAAAGGTACAGAGATTGTACTTCTATCAAAATTTTATGGAAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGGGTTCCTCCCAATTCAGCAGTTGTAGAAATGGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGGACCATCATGCTGTGATTAATTCAGATGCTCGATTTGGAAAAAGACTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAATTAATTAATAAAATTTTCTCTCTGGAATTAATAATAACAGATTTACTTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCCGGATATGTGAGGGCTTGCAATGTATGAGTGTAGAGAATGCTAC 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGGACATCTCAGCTGGAAAAATCAACGAGTTTTTGTAAAACTCGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTCATCCGAGAGGCTGAATCATATAATATAACCCAGTGTCTCTCCCAAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACACGCTGCATCCCTTCCAGAAATATGGAGTTATTGCT 2809
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAACCAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGAGTCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGCTCTCTTTGACAGCATGGCGGATCGGATGGTGTGAGATGGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyLysTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATCTCTGGACTCCAGGAGATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGATGTACCAGAGTCCAAATGAGTTTGTACAAA 3088

RESULT 10

US-09-397-424A-5019
; Sequence 5019, Application US/09397424A - *Abdu*
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN OSTEOBLAST LIBRARY
; FILE REFERENCE: MLN98-45PM
; CURRENT APPLICATION NUMBER: US/09/397,424A
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,469
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/106,454
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,252
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/132,100
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5379
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5019
; LENGTH: 4668

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
; US-09-397-424A-5019

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 21 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-397-424A-5019 (1-4668)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 230 ATGAGTTCTAGGCTTATGGAGCCCAAGAAAGATCACTTCCACCTTCTGGGAGAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
DB 290 TTTTACTTCTCTTCAAGAAATGACAGCTTTACAGACAAACAAACAAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 350 GTACCAAGAGGAGATAGACAGATATATTCAAGATCGTTCTGTGGGCAATTCAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnGlnProHisAla 80
DB 410 CCTTCTGCAAAAGGCAAGAAATACAGATTGGATTAAATTTAGAGCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
DB 470 GTTCTCTTTGTGATGAAAGGATGTTGTAGATATAATGAAAGTTTCACAGATTACTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
DB 530 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCAGCCTGTGTTTAAACAGAACACAGCTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 590 AAAGGCTCTCAATAGACGCTGGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 650 AATTTCTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGGACAGCTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
DB 710 ATATCTTTGGAGTTGAAATGCTGGAGAGGCTGCTGGTCAAGGTTTCACTGACGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 770 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTTCATTGGAC 829
QY 199 LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 830 AAGCTAGAACTCTAGAAGATGATGACACTGCAATTGGAAAGTGAATACGCAGGCTCTGGG 889
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 890 GACCAATGCGAGTCAACTTCTCTCTTGGAAATAACTCCAGAGTTCTTTTGAAGGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 950 GGAGAAACCAATAGAACTCGAACAGATTATATCTGTGATGTTTGGCAGGAAAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1010 TTAGGATATTTTGTGTGGACATGGATAACCTTATTGGCACTGGGATGGAAGATT 1069

```

```

QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1070 GATGGAGTGCAGCTTTGTAGTTTGGCTGTGTTGAAAGTACAAATCTATTGCCATCAAT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
DB 1130 GATATCATCCAGAGAGTGTGACGAGAAAGAGGCTCCCAAACCTTGCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1190 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCATATAAACCAAGGCTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1250 TCAGACCCTGGAAATGAAACAGATCTGAATATTTTATACCTTAAATGGGTCTTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1310 GACTCACACCAACCAATCCAAATCAAAAATACATGTTAGTTGAAGTTGCAGAGAC 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheArgSerProProLeuGln 396
DB 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTACCACCACTCCAG 1429
QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1430 CCTCCTCTGTGAACCTACACGACCCGAGAACAGATTCCACTCTTTACCATTGAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1490 ACCAAGATGCCCAATCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTAGCCGAG 1549
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1550 TCTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCATGCT 1609
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1610 CCTGGAAACTCACATGCTCTAGAAAGTGGGCTCATTTGGCTGAAAGTTAAGGAAACCTCCT 1669
QY 477 PheTyrGlyValIleArgTyrIleGlyGlnProGlyLeuAsnGluValLeuAlaGly 496
DB 1670 TTCTATGGGTAAATCCGTTGGATCGGTGCGTCCAGCCAGGACTGAAATGAAGTCTCGCTGA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1730 CTGGAACCTGGAAGATGATGCTGCGGCTGTACGATGGAACCTTCAGAGGCACTCGGTAT 1789
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1790 TTCACCTGTGCCCTGAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCGCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1850 AGTTTTCATCATTTGCAGCGGTTTCCAAATCAGATTGAGCGCTGTAATCTTCTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 1910 GGAAGGCTACTTAAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1970 TTGGAGATATGATTGGGAAGAGAAAGGCATCCAGGCTCATTAACATTTCTTTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2030 GACTCAACCTTATCTCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2090 CCCAAAGAAAGAACGATAGATATATTATAGTGAAACCCCAAGAGCTACTAGGAGACAGAA 2149

```

```
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATTGTTAACTCTGAGATATATGATATGTGTGCCCAAAAATATGAAACTGAG 2209
QY 657 LysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTGAAGGTGGGGCTGCATCAGGATTTACCTCTGAGAAAAGATCTCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTTGAATATCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTTAAATAAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAGATTTGTTACTTATCAAAATTTTATGGAATAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCGTCCCAATTCAGCAGTTGTTAGATGGTCTTTTATCAACAGTAACTCG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGGCACCATCATCTCTGATTAATTCAGTGCCTCGATTTGGAAGAAGCTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATTTTAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCCGGATATGTGAGGGCTTGCAATGATGAGTGTAGAGATGCTAC 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysValLysThrCysAsnThrGln 816
Db 2630 GAGCATCCGGACATCTCAGCTGGAAATATCAACAGTTTTTGTAAACCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGGCTGAATATAATATAAACCAGTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACAGCGCTGCATCCCTTGCCAGAATATGAGGATTAATTTGCT 2809
QY 857 ValLeuCysIleGlnThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGGAAGGACGATCT 2869
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCCTGGCTCTCTTTGACAGCATGGCCGATCGGATGGTGTGATGATGCTTCAACATT 2929
QY 897 ProGlnValThrProCysProGluValGlyGlyLysIleLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCAACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGACCTG 2989
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTTGACCTCCAGGAGAAATCCAAAGGCTGTGCAGAGATGCTTGTGTGATGATAT 3049
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGATGTACAGAGTCCCAACATGAGTTGTATCAAA 3088
```

RESULT 11

```
US-09-432-241A-3760
; Sequence 3760, Application US/09432241A - Abdn
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Hoitman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN EPIDERMAL KERATINOCYTE LIBRARY
; FILE REFERENCE: 1600.1004001
```

```
; CURRENT APPLICATION NUMBER: US/09/432,241A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/106,445
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,227
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/127,182
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/132,099
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5041
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3760
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-432-241A-3760
```

```
Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 22 Gaps: 4
```

US-09-671-687A-3 (1-949) X US-09-432-241A-3760 (1-4668)

```
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTTCAGCTTATGGAGCCCAAGAAAAGTCACTTCACTCCCTACTCGGAAGACGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCAAGAAATGCAGGTTACAGACAAACACAAACAAAGCTCTTAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCCGAAGGAAAGTAGTAGCAGATATATTCAAGATCGTTCTGTGGGCGCATTTCAAGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleGlnGlnProHisAla 80
Db 410 CCTTCTGCAAAAGCGCAAGAAAATCAGATTTGGATTTAAAAATTTAGAGCAACCTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTGTTCATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACTT 529
QY 100 LeuAlaIleThrAsnCysGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTAACCAATTTGTGAGAGAGGTTTCAGCGCTGTTTAAAAACAGAAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTTCAAAATAGACGTGGGCTGTCTGTGAAAAGTACAGCTGAGATCTGGGGGAAGAA 649
QY 140 LysPheProGlyValValArgPheArgPheGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTGGAGTTGTACGCTTTCAGAGGAGCCCTGTTTAGCAGAGACAGACAGCTCTCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTCTTGGAGTTGAATTTGCTGGAAGAGGTCGTGCTCAAGGTTTCACTGACGGGCTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTTGTGCGTGTGTTGTTGCAATGGAC 829
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
```

830 AGCTAGAACTCATAGAAATGATGACACTGCAATGGAAAGTGAATTTACGCAGGTCTCTGGG 889
219 AspThrMetGlnValGluLeuProProLeuGluLeuAAsnSerArgValSerLeuLysGly 238
890 GACACAATGCAGGTGAACTTCTCTCTTTGGAAATAAACTCCAGAGTTCTTTTGAAGTT 949
239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
950 GGAGAAACAATAAGAACTCTGGAACAGTATATCTGTGATGTTTGGCAGGAAAGAAGC 1009
259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
1010 TTAGGATATTTTGTGTGTGACATGGATTAACCTATTGGCAACTGGGATGGAGATTT 1069
279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
1070 GATGGAGTCAGCTTTGTAGTTTGTGCTGTGTTGAAAGTACAATTTCTATTGGACATCAAT 1129
298 AspIleIleProGluSerValThrGlnGluArgArgProProLysIleAlaPheMetSer 317
1130 GATATCATCCAGAGAGTGTGACGAGGAAAGAGGCCTCCCAAACTTGCCTTTATGTCA 1189
318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAACCAAGGCTACAGGATCTACC 1249
338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
1250 TCAGACCTCGGAAATAGAACAGATCTGAATATTTTATACCTTAATGGTCTTCTGTT 1309
357 AspSerGlnProGlnSerLysSerLysAsnThrTrpIleAspGluValAlaGluAsp 376
1310 GACTCACACCAACAATCCAAATCAAAATAATACATGTTGATGAAGTTGCAGAAGAC 1369
377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCTCCAG 1429
397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
1430 CTTCTCTCTGTGACTACTGACCCAGCAGAGACAGATTCACCTCTTTACCATTCAGTCTC 1489
417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
1490 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTCTGAGCCAG 1549
437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
1550 TCTGTAATGGAAGAGCTAAACACTGACCCGCTCCAAAGAGAGTCCACCCTTGGCCATGCCT 1609
457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
1610 CTTGGAACTCACATGCTTAGAAGTGGGCTCAATGGCTGAAGTTAAGGAGAACCTCTCT 1669
477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
1670 TTCTATGGGTAAATCCGTTGGATCGTTCAGCCACACGAGACTGATGAAGTGTCTCGTGG 1729
497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
1730 CTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTCAGAGGCACCTCGGTAT 1789
517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
1790 TTCACCTGTGCGCTGAAGAGGCGCTGTTGTGAAACTGAAGAGCTGCAGGCGCTGACTCT 1849
537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
1850 AGTTTGCATCATTTGCAGCGCGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576

1910 GGAGGCTACTTAAGTGAAGTAGAAGAAATACTCCACCACAAATAATCGAAAAAGAGGC 1969
577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
1970 TTGGAGATAATGATTGGAAAGAGAAAGGCATCCAGGCTCATTAACAATCTTGTGTACTTA 2029
597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
2030 GACTCAACCTTATTCTCTCTATTGCTTTTGTAGTTCTGTCTGGACACTGTGTACTTAGA 2089
617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
2090 CCCAAGAAAAAGAACGATGTAGATATATTAGTGAACCCCAAGAGCTACTGAGGACAGAA 2149
637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
2150 ATTTGTTAATCCTCTGAGAAATATATGATATGTGTGTGCACAAAAAATATGAAACTGAGG 2209
657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
2210 AAAATACTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAG 2269
677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
2270 GAATTTCTTGAATATTCGTCTTTCATCATATTTTAAAGGTAGAACCTTCTGTAAAAATAAGA 2329
697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
2330 TCAGCAGGTCAAAAGGTACAGATTTGTACTTCTATCAAAATTTTATCGAAAAAATGAG 2389
717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
2390 AAAAGTTGGCGTTCCCAATTCAGCAGTTGTAGATGGTCTTTTATCAACAGATAACCTG 2449
737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
2450 AAAATTTGCAGAGGACCATCATGTCTGATTATTACAGATGCCTCGATTTGGAAAAAGACTTT 2509
757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
2510 AAATATTTAAAAAATTTTCTCTCTGGAATTAATAATACAGATTTTACTTGAAGAC 2569
777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
2570 ACTCCACAGACGTCCGATATGTGGAGGCTTGCATATGATGAGTGTAGAGATGCTAC 2629
797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
2630 GACGATCCGGACATCTCAGCTGGAAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2689
817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
2690 GTCCACCTTCATCCGAAGAGGCTGAATCATATAATATTAACCCAGTGTCTACTTCCCAAGAC 2749
837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
2750 TTACCCGACTGGGACTGGACACGCGTGCATCCCTTGGCCAGATATCGAGTTATTGCT 2809
857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
2810 GTTCTCTGCATAGAAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGACGATTTCT 2869
877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
2870 GCCTGGCTCTTCTTTGACAGATGGCGGATGGTGGTCAAGATGGCTTCAACAT 2929
897 ProGlnValThrProCysProGluValGlyGlyLysTyrLeuLysMetSerLeuGluAspLeu 916
2930 CCTCAAGTCAACCCATGCCAGAGTAGAGAGTACTTTGAAGATGTCTCTCGAGAGACCTG 2989
917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
2990 CATTTCTTGACTCCAGAGAAATCCAAGGCTGTGCAGAGACTGCTTTGTGTATCATAT 3049

QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 DB 3050 ATGTGATGTACAGAGTCCAAACATGAGTTGTACAAA 3088

RESULT 12

US-09-434-737-1278

; Sequence 1278, Application US/09434737 — *Abdn*
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Kingsbury, Gillian A.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
 ; TITLE OF INVENTION: HUMAN MIXED LYMPHOCYTE LIBRARY
 ; FILE REFERENCE: 1600.1067001
 ; CURRENT APPLICATION NUMBER: US/09/434,737
 ; CURRENT FILING DATE: 1999-11-05
 ; EARLIER APPLICATION NUMBER: 60/107,228
 ; EARLIER FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 1830
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1278
 ; LENGTH: 4668
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(4668)
 ; OTHER INFORMATION: n = A,T,C or G

US-09-434-737-1278

Alignment Scores:

Pred. No.: 0 Length: 4668
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 22 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-434-737-1278 (1-4668)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
 DB 230 ATGAGTTCAGGCTATGGAGCCCAAGAAAAGTCACCTCCCTTCTGGGAAGACGGATT 289
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 DB 290 TTTTACTTCTCTTCAAGATCCAGCGTTACAGACAAACAAACAAAGCTCCCTTAA 349
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 DB 350 GTACCGAAGGAGTATAGACAGTATATTCAGATCGTTCTCTGGGGCATTCAAGATT 409
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 DB 410 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAATTAATAATCTAGACCAACCTCATGCA 459
 QY 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeu 99
 DB 470 GTTCTCTTTGTTGATGAAAGGATGTTGAGAGATAAATGAAAGTTTACAGAGTTACTT 529
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 DB 530 TTGGCAATTACCAATTGTGGAGAGAGTTTCAGCTGTTTAAACAGAAACAGACTAAGT 589
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 DB 590 AAAGGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGAAGAA 649
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 DB 650 AAATTTTCTGGAGTTGTACGCTTTACAGAGGCCCTGTGTACAGAGAGGACAGTCTCGGA 709
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179

DB 710 ATATTTCTTTGGAGTTGAAATTCCTGGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 DB 770 TACCRAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 DB 830 AAGCTAGAACTCATAGAGATGATGACACTGCACTGGAAAGTGATTACGCGAGGTCCTGGG 889
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 DB 890 GACACAAATGCAAGTCGAACCTCTCTTGGAAATAAATCCAGAGTTCTTTGAAGGTT 949
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 DB 950 GGAGAAACCAATAGAACTCGAACAGATTATTTCTGTGATGTTTGGCAGGAAAGAAAGC 1009
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
 DB 1010 TTAGGATATTTTGTGTGGACATGGATAACCCCTATTGGCAACTGGGATGGAAGATT 1069
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 DB 1070 GATGGAGTCAGCTTTTGTAGTTTGTGTGTGTTGAAAGTACAAATTTCTATTGCACATCAAT 1129
 QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
 DB 1130 GATATCATCCAGAGAGTGTGCGAGGAAAGAGGCTCCCAAACTTGCCTTTATGTCA 1189
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 DB 1190 AGAGGTGTTGGGACCAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC 1249
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 DB 1250 TCAGACCTCTGGAATAAGAAACAGATCTGAATATTTTATACCTTAAATGGGCTCTCTGTT 1309
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
 DB 1310 GATCAACCAACCAATCCAATCAAAAATACATGGTACATTGATGAAGTTGCAAGAGAC 1369
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGln 396
 DB 1370 CCTGCAAAATCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACCTCCAG 1429
 QY 397 ProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 DB 1430 CCTCTCTCTGTAACCTCACTGACCCAGAACAGATTCACACTCTTTACCAATTCACTCTC 1489
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 DB 1490 ACCAAGATGCCCAATCAATGGAAGTATTTGGCCACAGTCCACTTCTCTGTACGCCAG 1549
 QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 DB 1550 TCTGTATGGAAGAGCTAAACACTGACCCCGTCCAAGAGAGTCCACCTTGGCCATGGCT 1609
 QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
 DB 1610 CCTGGGAACCTCACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT 1669
 QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
 DB 1670 TTCTATGGGTAAATCCGTGTGGATCGGTGAGCCACAGGACTGAATGAAGTCTCGCTGGA 1729
 QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
 DB 1730 CTGAACTGGAAGATGATGTGAGGCTGTACGGATGGAACCTTCAGAGGACCTTCGGTAT 1789
 QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536

Db 1790 TTCACCTGTGCTGAAGAGCGCTGTTTGTGAACCTGAAGAGCTGCAGCGCTGACTCT 1849
Qy 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
Db 1850 AGGTTTGCATCATTTGCAGCGGTTTCCAAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 1909
Qy 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576
Db 1910 GGAGGCTACTTAAAGTAGTAGAAGAAATACCTCCCAAAATGGAAAAAGAGGC 1969
Qy 577 LeuGluIleMetGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
Db 1970 TTGGAGATATGATTGGGAAGAGAGGATCCAGGGTCATTACAAATCTTGTACTTAA 2029
Qy 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
Db 2030 GACTCAACCTTATCTGCTTATTTGCTTTATTTAGTTCTGTCGACACTGTGTTACTTAA 2089
Qy 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
Db 2090 CCCAAAGAAAGAACCATGTAGAAATATTATAGTGAACCCCAAGAGCTACTGAGACAGAA 2149
Qy 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
Db 2150 ATGTTTAACTCTGAGATATATGATATGCTGTGCCACAAAATTTATGAACCTGAGG 2209
Qy 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
Db 2210 AAAATACTTTGAAAGGTGAGGCTGCATCAGGATTTTACCTCTCGAAGAAAAGATCTGAG 2269
Qy 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
Db 2270 GAATCTCTGAATATCTGTTTCAATCATATTTTAAAGGGTAGAACCTTTTGTAAAAATAAGA 2329
Qy 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
Db 2330 TCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTCAAAATTTTATGGAAAAAATGAG 2389
Qy 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
Db 2390 AAAGTTGGCTTCCCAATTCAGCAGTTGTTAGAAATGGTCTTTTATCAACAGTAACCTG 2449
Qy 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
Db 2450 AAATTTGCAGAGCCACCATCATGCTGATATTATTCAGATGCTCGATTTGGAAAAAGACTTT 2509
Qy 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
Db 2510 AAACATATTTAAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGAC 2569
Qy 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
Db 2570 ACTCCAGACAGTCCGGATATGTGAGGCTTGCATGTTATGAGTGTAGAGATGCTAC 2629
Qy 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2630 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCAA 2689
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2690 GTCCACCTTCATCCGAAGAGGCTGAATATAATAATAAACCCAGTGTCACTTCCCAAGAC 2749
Qy 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 2750 TTACCCGACTGGGACTGGAGACAGCGCTGCATCCCTTCCAGAAATATGGAGTTATTGCT 2809
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
Db 2810 GTTCTCTGCATAGAAACAGCCACTATGTGCTTTTGTGAAGTATGGGAAGGAGGATCT 2869
Qy 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2870 GCTGGCTCTTCTTTCACAGCATGGCCGATCGGGATGGTGGTCAAGATGCTTCAACATT 2929

Qy 897 ProGlnValThrProCysProGluValGlyGlyTyrLeuLysMetSerLeuGluAspLeu 916
Db 2930 CCTCAAGTCACCCATGCCAGAAAGTAGGAGACTTGAAGATGTCTCTGGAAGACCTG 2989
Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 2990 CATTCCTGGACTCCAGAGAAATCCAAAGGCTGTGCAGAAAGACTGCTTTGTGATGCATAT 3049
Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3050 ATGTGCATGTACCAGAGTCCAAACAATGAGTTGTACAAA 3088

RESULT 13

US-09-850-118-1278 - *Abdn*
; Sequence 1278, Application US/09850118
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN MIXED LYMPHOCYTE LIBRARY
; FILE REFERENCE: 1600.1067-002
; CURRENT APPLICATION NUMBER: US/09/850.118
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/107,228
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 09/434,737
; NUMBER OF SEQ ID NOS: 1830
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1278
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-850-118-1278

Alignment Scores:

Pred. No.: 0 Length: 4668
Score: 4983.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 36 Gaps: 4

US-09-671-687A-3 (1-949) x US-09-850-118-1278 (1-4668)

Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTTCAGCTTATGGAGCCCAAGAAAAAGTCACTTCCCTACTGGAGAGCGGATT 289
Qy 21 PheTyrIleLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCTCAAGAAATGCAGCGTTACAGACAAACAAACACAAAAGCTCTCTAAA 349
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCGAAGGAAGATAGGACAGATATATTCAAGATCGTTCTGTGGGCGATTCAGAGGATT 409
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTCTCAAAAGCAAGAAAAATCAGATTGGATTAAAAATTTAGACCACTCATGCA 469
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAGATTTCACAGAGTTACTT 529
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTACCAATTGTGAGGAGAGGTTTACGCTGTTTAAAAAACAAGAAACAGACTAAGT 589

QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 590 AAAGGCTCCAAATAGACGTGGGCTCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 650 AAATTTCTTGGAGTTGACGCTTCAGAGACCCCTGTAGCAGAGGACAGTCTCCGGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGlyValArgGlyGlnGlyPheThrAspGlyVal 179
DB 710 ATATTTCTTGGAGTTGAAATGCTGTGGAAGAGTCTGTGCTCAAGGTTTCACTGACGGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829
QY 199 LysLeuLeuLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 830 AAGCTAGAACTCATAGAAGATGATGACCTGCAATTGGAAAGTGAATTACGAGGTCCTGGG 889
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 890 GACACAAATCGAGTCTGAACTTCTCTCTTGGAAATAACTCCAGAGTTTCTTTGAAGGTT 949
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
DB 950 CGAGAAACAATAGAACTCGAATCTGAACTATATCTGTGATGTTTGGCCAGGAAAGAAAGC 1009
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
DB 1010 TTAGGATATTTGTTGGTGTGGACATGGGATAACCTTATTTGGCAACTGGGATGGAAGATT 1069
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
DB 1070 GATGGAGTCAGCTTTGTAGTTTGTGTTGGTGTGTTGAAAGTACAATCTATTGGCACAAT 1129
QY 298 AspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
DB 1130 GATATCATCCAGAGAGTGTGACGACGAGGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 1189
QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
DB 1190 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACC 1249
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
DB 1250 TCAGACCTTGGAAATAGAAACAGATCTGAATATTTTATACCTTAAATGGGCTCTCTGTT 1309
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
DB 1310 GACTCACACCAATCCAAATCAAAAATACATGTTGATGATGATGATGATGATGATGATGAT 1369
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGln 396
DB 1370 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTCAACCACCTCCAG 1429
QY 397 ProProProValAsnSerIleThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
DB 1430 CCTCTCTGTGAACTCAGTACGACCCGAGAACAGATTCACCTCTTTTACCATTCAGTCTC 1489
QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
DB 1490 ACCAAGATGCCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTCTCTGTGAGCCGAC 1549
QY 437 SerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
DB 1550 TCTGTAATGGAAGAGCTAAACACTGACCCGCTCCAAAGAGAGTCCACCTTTGGCCATGCCT 1609
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 1610 CCTGGGAACCTCACATGCTTAGAAGTGGGCTCATTTGCTGAAAGTGAAGGAGAAACCTCTCT 1669

QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 1670 TTTCTATGGGGTAAATCCGTTGGATCGGTTCAGCCACAGGACTGAATGAAGTCTCGCTGGA 1729
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 1730 CTGGAACTGGAAGATGAGTGTGAGGCTGTACGATGGAACCTTTCAGAGGACTCGGTAT 1789
QY 517 PheThrCysAlaLeuLeuLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 1790 TTCACCTGTGCTGAAGAAGCGCTGTTGTGAACCTGAAGAGCTGCAGGCCCTGACTCT 1849
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 1850 AGGTTTGCATCATTCGAGCGGTTTCCAATCAGATTCAGCGCTGTAACTCTTTAGCATTT 1909
QY 557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
DB 1910 GGAGGCTACTTAAGTGAAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGGC 1969
QY 577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 1970 TTGGAGATTAATGATTGGGAAGAAAGCATCCAGGTCATTACAACTCTTGTACTTA 2029
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 2030 GACTCAACTTATCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTTACTTAGA 2089
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 2090 CCCAAAGAAAGAACCATGTAGAAATATATAGTGAACCCCAAGAGCTACTGAGGACAGAA 2149
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 2150 ATGTTAACTCTCGAATATATATGATATGATGATGATGATGATGATGATGATGATGAT 2209
QY 657 LysIleLeuGluLysValGluAlaIleSerGlyPheThrSerGluGluLysAspProGlu 676
DB 2210 AAAATCTTTGAAAGGTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAAAGATCTCTGAG 2269
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
DB 2270 GAATCTTGAATATCTGTTTTCATCATATATTTAAGGGTAGAACCTTTTGTCTAAATAAGA 2329
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 2330 TCAGCAGGTCAAAGGTACAAAGATTGTACTCTCTCAAAATTTTTATGGAAAAAATGAG 2389
QY 717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 2390 AAAGTTGGCGTCCCACTTCCAGCAGTTGTAGATGGTCTTTTATCAACAGTAACCTG 2449
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
DB 2450 AAATTTGCAGAGGACCATCATCTGATTTATTCAGATGCTCGATTTGGAAAAAGACTTT 2509
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 2510 AAATCTTTTAAAAAATTTTTCTCTCTGGAATTAATAATATAACAGATTTACTTTGAAGAC 2569
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 2570 ACTCCAGACAGTCCGGATATGTGGAGGCTTGGCAATGATGATGATGATGATGATGATGAT 2629
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 2630 GACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAACTCTGCAACACTCAA 2689
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 2690 GTCCACCTTCATCCGAGAGGCTGAATCATAAATATAAACCCAGGTGTCACTTCCCAAGAC 2749
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856

Db 2750 TTACCGAGCTGGAGTGGAGACACGGCTGCATCCCTTGCAGATATGAGTTATTGCT 2809
 QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
 Db 2810 GTTCTCTGCATAGAACCAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTCT 2869
 QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyClnAsnGlyPheAsnIle 896
 Db 2870 GCCTGGCTCTCTTTGACAGCATGGCCGATGGGATGGTGGTTCAGAAATGGCTTCAACATT 2929
 QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
 Db 2930 CCTCAAGTCACCCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTGGGAAGACCTG 2989
 QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
 Db 2990 CATTTCTTGACTCCAGGAGATCCAAAGCTGTGCAGGAGACTGCTTGTGTGATGCATAT 3049
 QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
 Db 3050 ATGTGCATGTACAGAGTCCAAATGAGTTGTACAAA 3088

RESULT 14

PCT-US02-27777-49
 ; Sequence 49, Application PC/TUS0227777
 ; GENERAL INFORMATION:
 ; APPLICANT: diaDexus, Inc.
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Salceda, Susana
 ; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
 ; FILE REFERENCE: DEX-0346
 ; CURRENT APPLICATION NUMBER: PCT/US02/27777
 ; CURRENT FILING DATE: 2002-08-29
 ; PRIOR APPLICATION NUMBER: US 60/316,306
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 170
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 49
 ; LENGTH: 6831
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 PCT-US02-27777-49

} too late

Alignment Scores:
 Pred. No.: 0 Length: 6831
 Score: 4983.00 Matches: 948
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 98.99% Indels: 4
 DB: 1 Gaps: 4

US-09-671-687A-3 (1-949) x PCT-US02-27777-49 (1-6831)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
 Db 1585 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCCCTTACTGGGAAGAGCGGATT 1644
 QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
 Db 1645 TTTTACTTCTCTTCAAGATGTCAGCGTGTACAGACAAACACAAAGCTCCTTAAA 1704
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
 Db 1705 GTACCAAGGGAAGTATAGACAGTATATTCAAGATCGTTCTGTGGGCAATTCAGGATT 1764
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
 Db 1765 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAATAATTCAGACCAACCTCATGCA 1824
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99

Db 1825 GTTCTCTTTTGTGATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTT 1884
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
 Db 1885 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCCCTGTTTAAAAACAGAAAACAGACTAAGT 1944
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
 Db 1945 AAAGGCTTCCAAATAGACGTGGGCTGTCCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 2004
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
 Db 2005 AAATTTCTGGAGTGTACGCTTCAGAGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 2064
 QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
 Db 2065 ATATTCTTTGGAGTTGAATTTGCTGGAAGAGGTCGTGCTCAAGGTTTTCACCTGACGGGGTG 2124
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
 Db 2125 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTTCATTGGAC 2184
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
 Db 2185 AAGCTAGAACTCATAGAGATGATGACACTGCATTGGAAAGTATTACCGAGGCTCTGGG 2244
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
 Db 2245 GACACAATGCAGGTCGAACCTTCCTCTTGGAAATAAATCCACAGATTTCCTTTGAAGGTT 2304
 QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
 Db 2305 GGAGAAACAATAGAACTCTGGAACAGATTATATCTGTGATGTTTTCGCCAGGAAAAAGAACG 2364
 QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspArgPhe 278
 Db 2365 TTAGGATATTTTGTGTGGTGGACATGATAACCTATTGGCAACTGGGATGGGAAGATT 2424
 QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
 Db 2425 GATGGAGTCAGCTTTGTAGTTTGTGTGTTGAAAGTACAATTCTATTGCACATCAAT 2484
 QY 298 AspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSer 317
 Db 2485 GATATCATCCAGAGAGTGTGACGAGAAAGAGGCTCCCAAACCTTCCTTTATGTCA 2544
 QY 318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
 Db 2545 AGAGGTGTTGGGGACAAAGGTTTCATCCAGTCAATAATAAACCAAGGCTACAGGATCTACC 2604
 QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356
 Db 2605 TCAGACCCCTGGAATAGAAACAGATCTGAATATTATTTATACCTTAAATGGGTCTTCTGTT 2664
 QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAsp 376
 Db 2665 GATTCACACCAACCAATCCAAATCAAAAAATACATGGTACATTGATGAAGTTGCAGAGAC 2724
 QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
 Db 2725 CCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTCACCGTTCTTCACCACCACTCCAG 2784
 QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
 Db 2785 CCTCTCTCTGTGAACCTCACTGACCACCGAGAACAGATTTCACACTCTTTTACCAATTGAGTCTC 2844
 QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436
 Db 2845 ACCAAGATGCCCAATCAATGGAAGTATTGGCCACAGTCCACCTTCTGTGTACGCCAG 2904
 QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
 Db 2905 TCTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCCATGCT 2964

QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
DB 2965 CTGGGAACCTCAGATGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCT 3024
QY 477 PheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
DB 3025 TTCATAGGGGTAAATCCGTTGGATCGGTGAGCCAGCCAGGACTGAATGAAGTGCCTCGCTGA 3084
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
DB 3085 CTGGAACTGGAGATGAGTGTGAGGCTGTACGGATGGAACTTTCAGAGGCACTCGGTAT 3144
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536
DB 3145 TTCACCTGTGCCCTGAAGAAGGCGCTGTTGTGAATCTGAAGAGCTGCAGGCGCTGACTCT 3204
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
DB 3205 AGTTTTCATCATTTGAGCGCGGTTTCCATCATGATGAGCGCTGTAACTCTTTAGCATTT 3264
QY 557 GlyGlyTyrLeuSerGluValGluGluAsnThrProProLysMetGluLysGluGly 576
DB 3265 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAGAGGC 3324
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
DB 3325 TTGGAGATAATGATTGGGAAGAGAAAGGCAATCCAGGGTCATTACAACTCTTGTACTTA 3384
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
DB 3385 GACTCAACCTTAATCTGCTATTGCTTTTGTGTTCTGTGACACTGTGTACTTAGA 3444
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
DB 3445 CCCAAAGAAAAGAACCATGTAGAAATATATAGTGAACCCAGAGCTACTGAGACAGAA 3504
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
DB 3505 ATTGTTAATCTCTGAGATATATGATGATGCTGTGCCACAAAATATGAACTGAGG 3564
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
DB 3565 AAAATACITTTGAAAGTGGAGGCTGATCAGGATTTTACCTCTCAAGAAAAGATCTGAG 3624
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleArg 696
DB 3625 GAATCTTGAATATCTGTTTTCATCATATATTTAAGGGTGAACCTTTGCTAAAAAATAGA 3684
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
DB 3685 TCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAATTTTATGGAAAAAATGAG 3744
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
DB 3745 AAAGTTGGCGTCCCAATTCAGCAGTGTGTAAGATGGTCTTTTATCAACAGTAACTG 3804
QY 737 LysPheAlaGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPhe 756
DB 3805 AAATTTGCAAGGACCATCATCTCTGATTATTATGATGCTCGATTTGGAAAAGACTTT 3864
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 3865 AAATTAATTAATAAATTTTCTCTCTGGAATTAATAATAACAGATTTACTTTGAAGAC 3924
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 3925 ACTCCAGACAGTCCGATATGTGGAGGCTTGTGCAATGTATGATGTAGAGATGTCTAC 3984
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 3985 GACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAAACCTGCAACACTCA 4044

QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 4045 GTCCACCTTTCTCCGAAGAGGCTGAATCATATAATAACCCAGTGTCACTTCCCAAAGAC 4104
QY 837 LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 4105 TTACCCGACTGGAGCTGGAGACACGGCTGCATCCCTTCCAGAAATATGGAGTTATTGCT 4164
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer 876
DB 4165 GTTCTCTGCATAGAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGAGCATCT 4224
QY 877 AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 4225 GCCTGCTCTCTTTTACAGCATGGCCGATCGGATCGGTGCTCAGAAATGGCTTCAACATT 4284
QY 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
DB 4285 CCTCAAGTCACCCCATGCCAGAAAGTAGGAGAGTACTTTGAAGATGTCCTCGGAAGACCTG 4344
QY 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
DB 4345 CATTCCTTGGACTCCAGGAGATCCAAAGCTGTGCAGAGACTGCTTTGTGATGCATAT 4404
QY 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 4405 ATGTGCATGTACCAGAGTCCAAATGAGTTTGTACAAA 4443
RESULT 15
PCT-US02-27777A-49
; Sequence 49, Application PC/TUS022777A
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Saiceda, Susana
; TITLE OF INVENTION: Compositions and Methods related to Breast Specific Genes and Proteins
; FILE REFERENCE: DEX-0346
; CURRENT APPLICATION NUMBER: PCT/US02/2777A } too late
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 60/316,307
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 6831
; TYPE: DNA
; ORGANISM: Homo sapien
PCT-US02-27777A-49
Alignment Scores:
Pred. No.: 0 Length: 6831
Score: 4993.00 Matches: 948
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 98.99% Indels: 4
DB: 1 Gaps: 4
US-09-671-687A-3 (1-949) x PCT-US02-27777A-49 (1-6831)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 1585 ATGAGTTTCAGCTTATGGAGCCCAAGAAAAGTCACTTCCCTACTTGGGAAGACGGATT 1644
QY 21 PheTyrLeuLeuGluGlnCysSerValThrAspLysGlnThrGlnLysLeuLys 40
DB 1645 TTTTACTTGTCTTCAAGAAATGACGCTTACACACAAACACAAAGCTCCTTAAA 1704
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 1705 GTACCCGAAGGAGTAGTAGACAGTATATTCAAGATCGTTCTGTGGGCAATTCAGGATT 1764
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80

1765 CTTCTGCAAGGCAAGAAATCAGATTGGATTAAAAATCTAGAGCAACCTCATGCA 1824
81 ValLeuPheValAspGlu---AspValValGluLeuAsnGluLysPheThrGluLeuLeu 99
1825 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 1884
100 LeuAlaIleThrAsnGluLeuArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
1885 TTGGCAATTTACCAATTTGTGAGAGAGGTTTCAGCCTGTTTAAAAACAAGAACACAGACTAAGT 1944
120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
1945 AAAGGCTCCAAATAGACGTGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 2004
140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
2005 AAATTTCTTGGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGAGGACAGTCTCCGGA 2064
160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
2065 ATATTCTTTGGAGTTGAAATGCTGGAGAGGCTGCTGGTCAAGGTTTCACTGACGGGGTG 2124
180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
2125 TACCAAGGAAACAGCTTTTTCAGTGTGATGAGATTTGGCGCTGTTGTTCATTGGAC 2184
199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
2185 AAGCTAGAACTCATAGAAAGATGATGACTGCAATGGAAGTGTACGAGGCTCTGGG 2244
219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
2245 GACACAAATCGAGTCTGAACTTCTCTCTTTGGAAATAAACTCCAGAGTTCTTTTGAAGTT 2304
239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258
2305 GGAGAAACAATAGAAATCTGGAACAGTTATATCTGTGATGTTTGGCAGGAAAGAAAGC 2364
259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPhe 278
2365 TTAGGATATTTTGTGTGTGACATGGAATTAACCTTATGGCAACTGGGATGGAAGATTT 2424
279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297
2425 GATGGAGTCAGCTTTGTAGTTTGGCTGTGTGAAAGTACAATTTCTATTGCAATCAAT 2484
298 AspileIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317
2485 GATATCATCCAGAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGCCTTTATGTCA 2544
318 ArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThr 337
2545 AGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAAATAAACCAAGGCTACAGGATCTACC 2604
338 SerAspProGlnSerLysSerLysAsnThrTyrIleAsnGlnGlySerSerVal 356
2605 TCAGACCCCTGGAAATAGAAACAGATCTGAATTAATTTTATACCTTAATGGGCTTCTGTT 2664
357 AspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAsp 376
2665 GACTCACACCAATCCAAATCAAAAATACATGTTGATGATGATGATGATGATGATGATGAT 2724
377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGln 396
2725 CTTGCAAAATCTTTACAGAGATATCTACAGACTTTTGGCCGTTCTTCCACCACCTCCAG 2784
397 ProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416
2785 CCTCTCTCTGTGAACACTCACTGACCCAGGAGAACAGATTTCCACTCTTTTACCATTCAGTCTC 2844
417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuAlaGln 436

2845 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG 2904
437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456
2905 TCTGTAATGGAAGAGCTAAACACTGCACCGGTCAAGAGAGTCCACCTTGGCCATGCT 2964
457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476
2965 CCTGGGAACCTCACATGCTTAGAAGTGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCT 3024
477 PheTyrGlyValIleArgTrrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496
3025 TTCTATGGGTGTAATCCGTTCGATCGGTTCAGCCACAGGACTGAATGAAGTGTCTGCTGGA 3084
497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516
3085 CTGGAACCTGGAAGATGATGTGAGGCTGTACGATGGAACCTTCAGAGGACCTCGGTAT 3144
517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspSer 536
3145 TTCACCTGTCCCTGGAAGAGGCGCTGTTGTGAACTGGAAGAGCTGCAGGCTGACTCT 3204
537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556
3205 AGTTTGCATCATTGACCGGTTTCCATCAGATTGAGCGCTGTAACTCTTTAGCATTT 3264
557 GlyGlyTyrLeuSerGluValValGluAsnThrProProLysMetGluLysGluGly 576
3265 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAGGC 3324
577 LeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596
3325 TTGAGATAATGATTGGGAAGAAAGGAGCATCCAGGCTCATTAACAATCTTGTGTACTTA 3384
597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616
3385 GACTCAACCTTATTTCTGCTTATTTGCTTTAGTTCTGTCTGGACACTGTGTACTTGA 3444
617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu 636
3445 CCCAAGAAAGAACGATGTAGATATATAGTGAACCCCAAGAGCTACTGAGGACAGAA 3504
637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656
3505 ATTGTTAATCTCTGAGATATATGATGTGTGTGTCACAAAAATATGAAACTGAGG 3564
657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676
3565 AAAATATCTTGAAGGTTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAG 3624
677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696
3625 GAATTTCTGAATATCTCTGTTTCATCATATTTAAGGTTAGAACCTTTCTTAAAAATAAGA 3684
697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716
3685 TCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGAAAAAATGAG 3744
717 LysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeu 736
3745 AAAGTTGGCGTTCCCAATTCAGCAGTTGTAGAAATGGTCTTTTATCAACAGATAACCTG 3804
737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
3805 AAATTTTCAGAGGACCATCATGTCTGATTATTTCAGATGCTCGATTTTGGAAAAAGACTTT 3864
757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
3865 AAATATTTAAAAAAATTTTCTCTCTGGAATTAATAATAACAGATTTACTTGAAGAC 3924
777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
3925 ACTCCACAGAGTCCCGGATATGTGGAGGCTTCAATGTATGATGATGATGATGATGATGATGAT 3984

QY	797	AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln	816
Db	3985	GACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTTGTAAAAACCTGCAACACTCAA	4044
QY	817	ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp	836
Db	4045	GTCCACCTTCATCCGNAAGGCTGAATCATATAATATAACCCAGTGTCACCTCCCAAAGAC	4104
QY	837	LeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla	856
Db	4105	TTACCCGACTGGGACTGGAGACACGCGCTGCATCCCTTGCAGAAATATGGAGTTATTGCT	4164
QY	857	ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSer	876
Db	4165	GTTCTCTGCATGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTCT	4224
QY	877	AlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle	896
Db	4225	GCCTGGCTCTTCTTTGACAGCATGGCCGATCGGGATGGTGGTCAGAAATGGCTTCAACATT	4284
QY	897	ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu	916
Db	4285	CCTCAAGTCACCCCATGCCAGAGTAGAGAGTACTTGAAGATGCTCTCTGGAGACCTG	4344
QY	917	HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr	936
Db	4345	CATTCTTGGACTCCAGGAGATCCAAAGGCTGTGCACGAAGACTGCTTTGTGATGCATAT	4404
QY	937	MetCysMetTyrGlnSerProThrMetSerLeuTyrLys	949
Db	4405	ATGTGCATGTACCAGAGTCCACAATGAGTTGTACAAA	4443

Search completed: April 18, 2005, 22:14:04
Job time : 7268 secs